PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Basic Commutators

M. A. Ward

Phil. Trans. R. Soc. Lond. A 1969 **264**, 343-412
doi: 10.1098/rsta.1969.0032

BASIC COMMUTATORS

By M. A. WARD

The Australian National University, Canberra

(*Communicated by B. H. Neumann, F.R.S.—Received* 18 *December* 1967)

CONTENTS

A theory of basic commutators is developed here which applies to polynilpotent series and related subgroups of a free group in much the same way as the conventional theory applies to the lower central series. The intention is to provide a theory of general application, however in §§ 17 and 19 new group-theoretic theorems are proved which involve the entire theory and illustrate its uses.

Introduction†

The conventional theory of basic commutators may be considered to be an investigation of the properties of the lower central series $\gamma_c(F): c = 1, 2, \ldots$ of an absolutely free group F, or alternatively of the properties of the corresponding factor groups, which are the free nilpotent groups of various classes:

$$F/\gamma_{c+1}(F) = F(\mathfrak{N}_c).$$

Suppose that G is a group generated by a subset $\mathscr{G}$. Then a set of 'formal expressions' may be constructed by using the elements of $\mathscr{G}$, the symbol 1 and the operations of inversion,

† Throughout this paper certain symbols in more or less common use will be employed without formal definition. For the reader's convenience these are collected in appendix II, together with all terms and symbols defined in the text.

multiplication and commutation. Each of these formal expressions will then represent a unique element of G in the obvious way. Certain formal expressions, known as *basic commutators of weight c* are defined for positive integers c and well-ordered by recursion over c as follows:

(i) The basic commutators of weight 1 are the elements of $\mathscr{G}$. They may be well-ordered in any way.

(ii) Assuming that $c > 1$ and that the basic commutators of weight $< c$ have been defined and ordered, the basic commutators of weight c are expressions of the form $[x, y]$ where x and y are basic commutators of weights r and s respectively, $r+s = c$, $x > y$ and, if $x = [x_1, x_2]$, then $x_2 \leqslant y$. The well-order may be extended to the basic commutators of weight c in any way so that they follow the commutators of smaller weight.

The following facts have been established:

Fact 1. A *collecting process* is defined, by means of which a formal expression for an element in $F(\mathfrak{N}_c)$, where $\mathscr{G}$ is a free generating set for this group, can be transformed into a particular type of expression known as a *basic product* of the form 1 or $b_1^{\beta_1} b_2^{\beta_2} \dots b_k^{\beta_k}$ where $b_1, b_2, \dots, b_k$ are basic commutators of weight $\leqslant c$, $b_1 < b_2 < \dots < b_k$ and $\beta_1, \beta_2, \dots, \beta_k$ are non-zero integers. The basic product represents the same element of the group as did the original expression.

Fact 2. The representation of a particular element of $F(\mathfrak{N}_c)$ in this form is unique.

Fact 3. $\gamma_c(F)/\gamma_{c+1}(F) \cong \gamma_c(F(\mathfrak{N}_c))$ is a free Abelian group, for which those elements represented by basic commutators of weight c constitute a free basis.

Fact 4. The upper and lower central series of $F(\mathfrak{N}_c)$ coincide. More specifically, provided the rank of F is > 1,

$$\zeta_r(F(\mathfrak{N}_c)) = \gamma_{c-r+1}(F(\mathfrak{N}_c)).$$

Fact 5. The lower central series of the absolutely free group F has trivial intersection:

$$\bigcap_{i=1}^{\infty} \gamma_i(F) = \{1\}.$$

In other words, F is residually nilpotent.

Fact 6. *Witt's formula*. When the number τ of generators is finite, the number of basic commutators of weight c is also finite and is the number

$$\frac{1}{c} \sum_{r|c} \mu(r)\, \tau^{c/r},$$

where μ is the Möbius function, defined for any positive integer r: $\mu(r) = 0$ if there exists $p > 1$ such that $p^2 | r$, $\mu(r) = (-1)^s$ otherwise, where s is the number of distinct primes dividing r.

The history of the subject can be covered briefly. The theory was initiated by P. Hall (1934) in a paper concerned with p-groups. Here the notion of basic commutator was introduced and the collecting process investigated (fact 1), however, the question of uniqueness was not treated in this paper. Witt (1937) showed that the whole question could be converted into an equivalent one concerning free Lie rings, and also produced the Witt formula (fact 6). Magnus (1935, 1937), also working in terms of free Lie rings, introduced the so-called *Magnus Ring* in terms of which the residual nilpotence of absolutely free groups (fact 5) was proved. Finally, M. Hall Jr (1950) proved the Basis Theorem (facts 2 and 3).

The work reported in this paper arose originally from the desire to prove the results of § 19 and more generally from the feeling that it should be possible to modify the theory of basic commutators as just described to permit the properties of free polynilpotent groups to be studied in the same way.

The idea of *weight* of a commutator is well known. This may be extended to the idea of weight of an expression (definition 1·3) and then the terms of the lower central series of a group G may be defined thus: an element of G belongs to $\gamma_c(G)$ if and only if it may be written as an expression of weight $\geqslant c$. This is possibly not a familiar way of defining the lower central series, and is made precise in definition 3·2 and lemma 5·2.

Here the idea of weight is generalized to that of *shape* and *shape range*. Then, for a given shape range W, shape subgroups $W_\alpha(G)$ of a group G, consisting of all elements of shape $\geqslant \alpha$ are defined. The generalization of *weight* to *shape* is the crux of this theory: for, just as the lower central series may be defined in terms of weight and then the conventional theory of basic commutators investigates the properties of this series, so the subgroups $W_\alpha(G)$ are defined in terms of shape and the theory to be described here investigates the properties of these subgroups. But the shape range W may be chosen so that these subgroups contain among them the terms of polycentral series. While the shape range W may be chosen so that subgroups not directly connected with polycentral series may be investigated, so that the early part of this paper will be slightly more general than the abstract suggests, the prime consideration throughout will be the study of polynilpotent groups.

In chapter I the idea of a shape range W and its associated shape σ are introduced. In terms of this the basic commutators are then defined. The most important part of this chapter consists of a proof that the number of basic commutators of a given weight, for a given number of generators, is independent of the shape range chosen to define them.

Chapter II develops the collecting process which, together with the main result of chapter I, provides the basis theorems. In this chapter the foregoing theory is applied to Lie rings and, as might be expected, the results are pleasingly straightforward.

In chapter III the shapes used to investigate the properties of free polynilpotent groups, the *polyweights*, are defined and an important result, namely that the corresponding factor groups are residually nilpotent, is proved.

In chapter IV the centralizers of shape subgroups modulo others are calculated. This allows the calculation of the upper central series of free groups of the varieties $\mathfrak{P}$, $\mathfrak{P} \wedge \mathfrak{N}_c$ and $\mathfrak{P} \vee \mathfrak{N}_c$ where $\mathfrak{P}$ is any polynilpotent variety.

During the development of the theory the results analogous to facts 1 to 6 above will be stated and proved.

While this paper was being prepared, Gorčakov (1967) published some related results. He defines multinilpotent (given in the translated summary as multipolynilpotent) varieties to be those formed from the nilpotent varieties by a finite number of applications of the operations of multiplication and intersection and gives the following theorem: Let A be a free group of some multinilpotent variety. Then

$$\bigcap_{n=1}^{\infty} \gamma_n(A) = \{1\}$$

and $\gamma_n(A)/\gamma_{n+1}(A)$ is a free Abelian group for each positive integer n. For the varieties defined in this paper by polyweights, these results are contained in theorems 15·1 and 9·1 (D) respectively.

The conventional theory of basic commutators concerns itself much of the time with formal expressions: not so much the elements of a group themselves as the way they are written down. In the present theory it will be found that more and more emphasis is placed on this aspect and that group-theoretic results, though the primary object of this study, appear infrequently. In order to avoid this essentially metamathematical approach, formal expressions for elements of a group are here replaced by elements of a free algebra. This algebra is chosen so as to be anarchic enough for us to regard (intuitively) the elements of the algebra as being in one-to-one correspondence with the possible formal expressions for elements in the group.

Chapter I. Basic commutators

1. *The algebra of expressions*

Definition 1·1. *Let* $\mathbf{G} = \{\mathbf{g}_i\}_{i<\tau}$ *be some set indexed by the ordinals less than some ordinal* τ, *the indexing being one-to-one. Form the algebra* $\mathbf{A} = \mathbf{A}_\Omega$ *generated freely by the set* $\mathbf{G}$ *with operator domain* $\Omega = \{\epsilon, \nu, \mu, \chi\}$ *where* ϵ *is a nullary operator* (*the* identity), ν *is a unary operator* (inversion), μ *and* χ *are binary operators* (multiplication *and* commutation *respectively*), *the only law being that* μ *is associative.*

A more conventional notation will be used for the effect of the operators on $\mathbf{A}$ *as follows*:

$$\epsilon = \mathbf{1},$$

$$\left.\begin{aligned} \mathbf{x}\nu &= \mathbf{x}^{-1} \\ \mathbf{x}\mathbf{y}\mu &= \mathbf{x}\mathbf{y} \\ \mathbf{x}\mathbf{y}\chi &= [\mathbf{x}, \mathbf{y}] \end{aligned}\right\} \textit{for all} \quad \mathbf{x}, \mathbf{y} \in \mathbf{A}.$$

Parentheses will be used in connexion with the operations of inversion and multiplication in accordance with the usual conventions. A 'left-normed' convention will be used in connexion with the operation of commutation, that is, $[\mathbf{a}, \mathbf{b}, \mathbf{c}] = [[\mathbf{a}, \mathbf{b}], \mathbf{c}]$ *and so on. The set* $\mathbf{G}$ *will be considered to be a subset of* $\mathbf{A}$ *in the usual way. The elements of* $\mathbf{A}$ *will be called* expressions *and* $\mathbf{A}$ *itself the* algebra of expressions.

It should be remarked that the operators ϵ and ν are not *bona fide* identity and inversion operators with respect to μ, since the associative law of multiplication is the only law of $\mathbf{A}$.

An elementary property of such an algebra is that, to each $\mathbf{a} \in \mathbf{A}$, there exists a uniquely defined positive integer which will here be called the *height* of $\mathbf{a}$ and denoted ht $(\mathbf{a})$, which is defined by its properties:

(i) $\mathrm{ht}(\mathbf{1}) = 1$ and if $\mathbf{g}_i \in \mathbf{G}$ then $\mathrm{ht}(\mathbf{g}_i) = 1$,

(ii) $\mathrm{ht}(\mathbf{a}^{-1}) = \mathrm{ht}(\mathbf{a}) + 1$, and

(iii) $\mathrm{ht}(\mathbf{ab}) = \mathrm{ht}([\mathbf{a}, \mathbf{b}]) = \mathrm{ht}(\mathbf{a}) + \mathrm{ht}(\mathbf{b}) + 1$.

Roughly speaking, the height of an expression is the number of symbols other than parentheses required to write it in terms of the generators $\mathbf{G}$ and the operators ϵ, ν, μ and χ.

It will be useful to have the following rather artificial definition of exponentiation of an expression.

Definition 1·2. *Let* n *be an integer and* $\mathbf{x} \in \mathbf{A}$. *Then* $\mathbf{x}^n$ *is defined recursively*:

(i) $\mathbf{x}^0 = \mathbf{1}$,

(ii) $\mathbf{x}^1 = \mathbf{x}$ *and* $\mathbf{x}^{-1}$ *is as described in definition* 1·1,

(iii) *for* $n > 1$, $\mathbf{x}^n = \mathbf{x}^{n-1}\mathbf{x}$ *and* $\mathbf{x}^{-n} = \mathbf{x}^{-(n-1)}\mathbf{x}^{-1}$.

Definition 1·3. *Let* N^- *denote the set of positive integers with an extra element* ∞ *adjoined and the usual addition and order extended to encompass* ∞ *by*

$$\left.\begin{array}{r}\infty+n = n+\infty = \infty+\infty = \infty\\ n<\infty\end{array}\right\}\quad \textit{for any positive integer } n.$$

The mapping wt: $\mathbf{A} \to N^-$ *is defined recursively over the height of expressions by*:

(i) $\mathrm{wt}\,(\mathbf{1}) = \infty$, $\mathbf{g}_i \in \mathbf{G} \Rightarrow \mathrm{wt}\,(\mathbf{g}_i) = 1$,

(ii) $\mathrm{wt}\,(\mathbf{x}^{-1}) = \mathrm{wt}\,(\mathbf{x})$,

(iii) $\mathrm{wt}\,(\mathbf{xy}) = \min\{\mathrm{wt}\,(\mathbf{x}), \mathrm{wt}\,(\mathbf{y})\}$, *and*

(iv) $\mathrm{wt}\,([\mathbf{x},\mathbf{y}]) = \mathrm{wt}\,(\mathbf{x}) + \mathrm{wt}\,(\mathbf{y})$.

For each expression $\mathbf{x}$, wt ($\mathbf{x}$) *will be called the* weight *of* $\mathbf{x}$.

Definition 1·4. *The closure (in* $\mathbf{A}$*) of* $\mathbf{G}$ *under the operation of commutation alone is the set* $\mathbf{C}$ *of* commutators.

For any expression $\mathbf{x} \in \mathbf{A}$, *a subset* $\Xi\,(\mathbf{x})$ *of* $\mathbf{C}$, *the* commutator set *of* $\mathbf{x}$, *is defined recursively by*:

(i) $\Xi(\mathbf{1}) = \varnothing$; $\mathbf{g}_i \in \mathbf{G} \Rightarrow \Xi(\mathbf{g}_i) = \{\mathbf{g}_i\}$.

(ii) $\Xi(\mathbf{x}^{-1}) = \Xi(\mathbf{x})$,

(iii) $\Xi(\mathbf{xy}) = \Xi(\mathbf{x}) \cup \Xi(\mathbf{y})$,

(iv) $\Xi([\mathbf{x},\mathbf{y}]) = \{[\mathbf{a},\mathbf{b}] : \mathbf{a} \in \Xi(\mathbf{x}),\ \mathbf{b} \in \Xi(\mathbf{y})\}$.

If $\mathscr{P}$ *is any property defined on the set of commutators and* $\mathbf{x}$ *is any expression in* $\mathbf{A}$ *then* $\mathbf{x}$ *has* $\mathscr{P}$ *essentially if every member of* $\Xi(\mathbf{x})$ *has* $\mathscr{P}$.

Clearly, if $\mathbf{c}$ is a commutator then either wt ($\mathbf{c}$) = 1 and $\mathbf{c} \in \mathbf{G}$ or else wt ($\mathbf{c}$) > 1 and there exist unique $\mathbf{x}, \mathbf{y} \in \mathbf{A}$, both of weight less than that of $\mathbf{c}$, such that $\mathbf{c} = [\mathbf{x},\mathbf{y}]$. A commutator has a property $\mathscr{P}$ essentially if and only if it has that property.

The emphasis throughout this study will be on groups, however some of the results will apply to a more general class of algebra. This may be defined as follows.

Definition 1·5. *A* describable algebra $G = G_\Omega$ *is one with the same operator domain* Ω *as* $\mathbf{A}$ *(definition* 1·1*) in which multiplication is associative. Except where otherwise stated the effects of the operators on* G *will be written in the usual way*: $\epsilon = 1$, $xv = x^{-1}$, $xy\mu = xy$ *and* $xy\chi = [x,y]$. *If* A *and* B *are two subalgebras of* G *then* $[A, B]$ *is the subalgebra generated by the set* $\{[a,b] : a \in A,\ b \in B\}$. *An* ideal *of* G *is a subalgebra* A *of* G *such that, if* $a \in A$ *and* $g \in G$, *then* $[a,g] \in A$ *and* $[g,a] \in A$.

The method whereby the elements of $\mathbf{A}$ describe the elements of such an algebra and so supplant the notion of 'formal expression' is now made precise.

Definition 1·6. *Let* G *be any describable algebra (or in particular a group) and suppose* G *is generated by a set* $\mathscr{G} = \{g_i\}_{i<\tau}$. *Let* $\mathbf{A}$ *be defined as in definition* 1·1. *Then the unique epimorphism* $\rho: \mathbf{A} \to G$ *for which* $\mathbf{g}_i\rho = g_i$ $(i < \tau)$ *will be called a* description *of* G; *moreover, if* G *is a relatively free algebra, freely generated by* $\mathscr{G}$, *then* ρ *will be called a* free description.

2. *Shape*

Definition 2·1. *A* shape range *is a quadruple* $(W, \leqslant, \preccurlyeq, +)$ *where* W *is a set,* $\leqslant$ and $\preccurlyeq$ *orders on* W *and* $+$ *a binary algebraic operation (addition) on* W *satisfying*:

(i) *The order* $\leqslant$ *(fully) well-orders* W. W *has a least element* 1 *and a greatest element* ∞. *The* *order* $\preccurlyeq$ *is a partial lattice order which implies* $\leqslant$ *in the sense that* $\alpha \preccurlyeq \beta \Rightarrow \alpha \leqslant \beta$. *The meet and join of* α *and* β *(with respect to* $\preccurlyeq$*) will be denoted* $\alpha \wedge \beta$ *and* $\alpha \vee \beta$ *respectively. The strict orders corresponding to* $\leqslant$ *and* $\preccurlyeq$ *will be written* $<$ *and* $\prec$ *respectively.*

(ii) W *is closed under addition and* $W - \{\infty\}$ *is generated by* 1 *under addition. Addition is commutative but not necessarily associative; a left-normed convention will be used:* $\alpha+\beta+\gamma = (\alpha+\beta)+\gamma$ *and so on.*

(iii) $\alpha, \beta \neq \infty \Rightarrow \alpha+\beta \neq \infty, \quad \alpha+\infty = \infty, \quad \alpha_1 < \alpha_2$ *and* $\beta \neq \infty \Rightarrow \alpha_1+\beta < \alpha_2+\beta$.

(iv) $\alpha \neq \infty \Rightarrow \alpha \prec \alpha+\beta, \quad \alpha_1 \prec \alpha_2$ *and* $\beta \neq \infty \Rightarrow \alpha_1+\beta \prec \alpha_2+\beta$.

(v) $\alpha \leqslant \beta \leqslant \gamma \Rightarrow \gamma+\beta+\alpha = \gamma+\alpha+\beta \preccurlyeq \beta+\alpha+\gamma$.

The orders $\leqslant$ *and* $\preccurlyeq$ *will be called the* fine *and* coarse *orders respectively.*

It will usually be possible to speak of the 'shape range W' rather than the less convenient 'shape range $(W, \leqslant, \preccurlyeq, +)$' without fear of confusion, for it will seldom be necessary to consider more than one shape range structure on one underlying set.

Lemma 2·1. *Let* $(W, \leqslant, \preccurlyeq, +)$ *be a shape range. Then so is* $(W, \leqslant, \leqslant, +)$.

Proof. It is sufficient to prove that the properties of $\preccurlyeq$ listed in definition 2·1 are also properties of $\leqslant$; that $\leqslant$ is a partial lattice order which implies $\leqslant$ is trivial, that

$$\alpha \neq \infty \Rightarrow \alpha < \alpha+\beta \quad \text{and} \quad \alpha \leqslant \beta \leqslant \gamma \Rightarrow \gamma+\beta+\alpha = \gamma+\alpha+\beta \leqslant \beta+\alpha+\gamma$$

follow from the corresponding properties for $\preccurlyeq$ and that

$$\alpha_1 < \alpha_2 \quad \text{and} \quad \beta \neq \infty \Rightarrow \alpha_1+\beta < \alpha_2+\beta$$

is stated explicitly in the definition.

This proof shows that the lemma can be stated in an alternative form: if, throughout the defining properties of a shape range, the coarse order is replaced by the fine one the resulting propositions are true. This in turn implies a metatheorem:

If $\mathscr{T}$ *is a theorem concerning a shape range* $(W, \leqslant, \preccurlyeq, +)$ *and this is translated into a new proposition* $\mathscr{T}'$ *by replacing* $\preccurlyeq$ *whenever it appears by* $\leqslant$ *and any expression defined in terms of* $\preccurlyeq$ *by one defined similarly in terms of* $\leqslant$ *then* $\mathscr{T}'$ *is a theorem (true proposition).*

Definition 2·2. *Let* Φ *and* Ψ *be two subsets of a shape range* W. *Then*

(i) $\Phi \preccurlyeq \Psi$ *if for every* $\psi \in \Psi$ *there exists* $\phi \in \Phi$ *such that* $\phi \preccurlyeq \psi$. *If further there exists* $\phi \in \Phi$ *such that* $\psi \in \Psi \Rightarrow \psi \not\preccurlyeq \phi$ *then* $\Phi \prec \Psi$.

(ii) $\Phi + \Psi$ *is the set* $\{\phi+\psi : \phi \in \Phi, \psi \in \Psi\}$.

(iii) $\Phi \vee \Psi$ *is the set* $\{\phi \vee \psi : \phi \in \Phi, \psi \in \Psi\}$.

Lemma 2·2. *With the notation of definition* 2·2,

(i) *The relation* $\preccurlyeq$ *pre-orders the subsets of* W. *The set* $\Phi \vee \Psi$ *has the property that, for any subset* Σ *of* W,

$$\Sigma \succcurlyeq \Phi \vee \Psi \Leftrightarrow \Sigma \succcurlyeq \Phi \quad \textit{and} \quad \Sigma \succcurlyeq \Psi$$

so the symbol $\vee$ *is appropriate. The relation* $\prec$ *on the subsets of* W *is the corresponding strict relation.*

(ii) *Addition of subsets of* W *is commutative but not necessarily associative.*

(iii) $\Phi + \varnothing = \varnothing$ *and* $\Phi + \{\infty\} = \{\infty\}$.

(iv) *If* $\Phi_1 \prec \Phi_2$ *and* $\Psi \not\succcurlyeq \{\infty\}$ *then* $\Phi_1 + \Psi \prec \Phi_2 + \Psi$.

(v) $\{\phi\} \preccurlyeq \{\psi\} \Leftrightarrow \phi \preccurlyeq \psi$, $\{\phi\}+\{\psi\} = \{\phi+\psi\}$ *and* $\{\phi\} \vee \{\psi\} = \{\phi \vee \psi\}$.

Proof. This all follows immediately from definitions 2·1 and 2·2, except perhaps for part (iv). Suppose then that $\Phi_1 \prec \Phi_2$ and $\Psi \not\geqslant \{\infty\}$. Then, for each $\phi_2 + \psi \in \Phi_2 + \Psi$ there exists $\phi_1 \in \Phi_1$ such that $\phi_1 \leqslant \phi_2$ and then $\phi_1 + \psi \in \Phi_1 + \Psi$ and $\phi_1 + \psi \leqslant \phi_2 + \psi$. This proves that $\Phi_1 + \Psi \leqslant \Phi_2 + \Psi$. But $\Phi_1 \prec \Phi_2$ so there exists $\phi_2 \in \Phi_2$ such that $\phi_1 \in \Phi_1 \Rightarrow \phi_1 \not\geqslant \phi_2$. There also exists $\psi \in \Psi$ such that $\psi \neq \infty$, so for any $\phi_1 + \psi \in \Phi_1 + \Psi$, $\phi_1 + \psi \not\geqslant \phi_2 + \psi$ by definition 2·1 (iv).

Definition 2·3. *Let* **A** *be an algebra of expressions and* W *a shape range. The* fine shape *on* **A** *associated with* W *is the mapping* σ: $\mathbf{A} \rightarrow W$ *defined by its properties*

(i) $\sigma(\mathbf{1}) = \infty$; $\mathbf{g}_i \in \mathbf{G} \Rightarrow \sigma(\mathbf{g}_i) = 1$.
(ii) $\sigma(\mathbf{x}^{-1}) = \sigma(\mathbf{x})$.
(iii) $\sigma(\mathbf{xy}) = \min\{\sigma(\mathbf{x}), \sigma(\mathbf{y})\}$, *where* min *is defined with respect to the fine order* $\leqslant$ *on* W.
(iv) $\sigma([\mathbf{x}, \mathbf{y}]) = \sigma(\mathbf{x}) + \sigma(\mathbf{y})$.

Corresponding to the coarse order is a coarse shape $\hat{\sigma}$: $\mathbf{A} \rightarrow W$ *defined by*:

(i)′ $\hat{\sigma}(\mathbf{1}) = \infty$; $\mathbf{g}_i \in \mathbf{G} \Rightarrow \hat{\sigma}(\mathbf{g}_i) = 1$.
(ii)′ $\hat{\sigma}(\mathbf{x}^{-1}) = \hat{\sigma}(\mathbf{x})$.
(iii)′ $\hat{\sigma}(\mathbf{xy}) = \hat{\sigma}(\mathbf{x}) \wedge \hat{\sigma}(\mathbf{y})$.
(iv)′ $\hat{\sigma}([\mathbf{x}, \mathbf{y}]) = \hat{\sigma}(\mathbf{x}) + \hat{\sigma}(\mathbf{y})$.

For each expression $\mathbf{x} \in \mathbf{A}$, *its* shape set $\Sigma(\mathbf{x}) \subseteq W$ *is defined by*

(i)″ $\Sigma(\mathbf{1}) = \varnothing$; $\mathbf{g}_i \in \mathbf{G} \Rightarrow \Sigma(\mathbf{g}_i) = \{1\}$.
(ii)″ $\Sigma(\mathbf{x}^{-1}) = \Sigma(\mathbf{x})$.
(iii)″ $\Sigma(\mathbf{xy}) = \Sigma(\mathbf{x}) \cup \Sigma(\mathbf{y})$.
(iv)″ $\Sigma([\mathbf{x}, \mathbf{y}]) = \Sigma(\mathbf{x}) + \Sigma(\mathbf{y})$.

Comparison of the three main sections of this definition shows that for any commutator **c**, $\sigma(\mathbf{c}) = \hat{\sigma}(\mathbf{c})$ and $\Sigma(\mathbf{c}) = \{\sigma(\mathbf{c})\}$. Further, comparison with definition 1·4 shows that, for any expression **x**, $\Sigma(\mathbf{x})$ could be defined alternatively by

$$\Sigma(\mathbf{x}) = \{\sigma(\mathbf{c}) : \mathbf{c} \in \Xi(\mathbf{x})\}.$$

A simple induction over the height of **x** shows that $\Sigma(\mathbf{x})$ is a finite set for any $\mathbf{x} \in \mathbf{A}$, that $\sigma(\mathbf{x})$ is the minimum member under the fine order $\leqslant$ of $\Sigma(\mathbf{x})$ and $\hat{\sigma}(\mathbf{x})$ is the meet of $\Sigma(\mathbf{x})$ and thus that $\hat{\sigma}(\mathbf{x}) \leqslant \sigma(\mathbf{x})$.

Another easy consequence of this definition is that N^- (definition 1·3) or more precisely $(N^-, \leqslant, \leqslant, +)$ is a shape range and that wt: $\mathbf{A} \rightarrow N^-$ is both the associated fine and coarse shape. This verifies the remark made in the introduction that 'shape' is a generalization of 'weight'. In view of results developed in the sequel it is of interest to observe that the only possible order $\leqslant$ for N^- for which $(N^-, \leqslant, \leqslant, +)$ is a shape range is in fact $\leqslant$ itself. When N^- is chosen as the shape range for any of the succeeding theorems it will be found that they reduce to known ones from the conventional theory of basic commutators or to trivialities. New group theoretic results must await the definition of *polyweights* in chapter III.

Both the fine and coarse shapes emerge as generalizations of weight and motivation for their definitions may be stated informally as follows: the axioms for the fine shape extract just those properties of weight necessary to make possible a definition of basic commutators and collecting processes which will allow the theory outlined in the introduction to be developed. The coarse order is chosen in such a way that it can utilize the basic commutators and collecting processes of the fine order to produce an analogous but new set of results. The justification for introducing the coarse order is that, in chapter III when a constructive definition of certain shapes is given, the subgroups corresponding to the coarse order turn out to be of more immediate general interest than those corresponding to the fine order and in consequence the group-theoretic results follow suit.

The development of analogous theories for the fine and coarse orders and the order defined on subsets in the sequel involves the statement of many theorems in three forms—one for each order. It will usually be possible to prove results only for the order on subsets, then infer the corresponding result for the coarse order by particularizing to subsets containing exactly one element (see lemma 2·2 (v)) and from this infer the corresponding result for the fine order by use of the metatheorem of § 2; this will avoid unnecessary triplication of proofs.

3. *Shape subgroups*

Definition 3·1. *Let* $\mathbf{A}$ *be an algebra of expressions,* $\sigma: \mathbf{A} \to W$ *and* $\hat{\sigma}: \mathbf{A} \to W$ *be a fine and coarse shape respectively and* Σ *the corresponding shape set function. Then for* $\alpha \in W$ *the sets* $\mathbf{W}_\alpha$ *and* $\hat{\mathbf{W}}_\alpha$ *are defined*

$$\mathbf{W}_\alpha = \{\mathbf{x}: \mathbf{x} \in \mathbf{A},\ \sigma(\mathbf{x}) \geqslant \alpha\}, \quad \hat{\mathbf{W}}_\alpha = \{\mathbf{x}: \mathbf{x} \in \mathbf{A},\ \hat{\sigma}(\mathbf{x}) \succcurlyeq \alpha\},$$

and for any subset Φ *of* W *the set* $\hat{\mathbf{W}}_\Phi$ *is defined*

$$\hat{\mathbf{W}}_\Phi = \{\mathbf{x}: \mathbf{x} \in \mathbf{A},\ \Sigma(\mathbf{x}) \succcurlyeq \Phi\}.$$

Lemma 3·1. *With the notation of definition* 3·1, *the sets* $\mathbf{W}_\alpha$, $\hat{\mathbf{W}}_\alpha$ *and* $\hat{\mathbf{W}}_\Phi$ *are fully invariant and hence verbal subalgebras of* $\mathbf{A}$.

Proof. For any endomorphism θ of $\mathbf{A}$ and $\mathbf{x} \in \mathbf{A}$,

$$\sigma(\mathbf{x}\theta) \geqslant \sigma(\mathbf{x}), \quad \hat{\sigma}(\mathbf{x}\theta) \succcurlyeq \hat{\sigma}(\mathbf{x}) \quad \text{and} \quad \Sigma(\mathbf{x}\theta) \succcurlyeq \Sigma(\mathbf{x}):$$

this is easily checked by induction over the height of $\mathbf{x}$. The lemma then follows immediately.

Definition 3·2. *Let* G *be a describable algebra (or in particular a group), let* $\rho: \mathbf{A} \to G$ *be a description of* G *and* σ, $\hat{\sigma}$ *and* Σ *be the shapes on* $\mathbf{A}$ *associated with a shape range* W. *Then for each* $\alpha \in W$ *the subsets* $W_\alpha(G)$ *and* $\hat{W}_\alpha(G)$ *of* G *are defined*

$$W_\alpha(G) = \mathbf{W}_\alpha \rho = \{\mathbf{x}\rho: \mathbf{x} \in \mathbf{A},\ \sigma(\mathbf{x}) \geqslant \alpha\},$$

$$\hat{W}_\alpha(G) = \hat{W}_\alpha \rho = \{\mathbf{x}\rho: \mathbf{x} \in \mathbf{A},\ \hat{\sigma}(\mathbf{x}) \succcurlyeq \alpha\},$$

and for any subset Φ *of* W *the subset* $\hat{W}_\Phi(G)$ *of* G *is defined*

$$\hat{W}_\Phi(G) = \hat{W}_\Phi \rho = \{\mathbf{x}\rho: \mathbf{x} \in \mathbf{A},\ \Sigma(\mathbf{x}) \succcurlyeq \Phi\}.$$

All these subsets will be called shape subalgebras *of* G. $\mathfrak{W}_\alpha$, $\hat{\mathfrak{W}}_\alpha$ and $\hat{\mathfrak{W}}_\Phi$ *are the classes of groups* G *for which* $W_\alpha(G)$, $\hat{W}_\alpha(G)$ *and* $\hat{W}_\Phi(G)$ *respectively are trivial: it follows from the next theorem that these are varieties.*

Theorem 3·1. *With the notation of the preceding definition,*

(i) $W_\alpha(G)$, $\hat{W}_\alpha(G)$ *and* $\hat{W}_\Phi(G)$ *are independent of the particular description* ρ *of* G *chosen to define them.*

(ii) *They are all verbal and hence fully invariant ideals of* G.

(iii) $W_1(G) = \hat{W}_1(G) = \hat{W}_{\{1\}}(G) = G$, $\quad W_\infty(G) = \hat{W}_\infty(G) = \hat{W}_{\{\infty\}}(G) = \hat{W}_\varnothing(G) = I$, *where* I *is the trivial ideal of* G, *that is, the ideal generated by* $\{1\}$.

(iv) $\alpha \leqslant \beta \Rightarrow W_\alpha(G) \supseteq W_\beta(G)$, $\alpha \preccurlyeq \beta \Rightarrow \hat{W}_\alpha(G) \supseteq \hat{W}_\beta(G)$, $\Phi \preccurlyeq \Psi \Rightarrow \hat{W}_\Phi(G) \supseteq \hat{W}_\Psi(G)$ *and as a special case of this last implication*

$$\Phi \subseteq \Psi \Rightarrow \hat{W}_\Phi(G) \subseteq \hat{W}_\Psi(G).$$

(v) $[W_\alpha(G), W_\beta(G)] \subseteq W_{\alpha+\beta}(G)$, $[\hat{W}_\alpha(G), \hat{W}_\beta(G)] \subseteq \hat{W}_{\alpha+\beta}(G)$,

$$[\hat{W}_\Phi(G), \hat{W}_\Psi(G))] \subseteq \hat{W}_{\Phi+\Psi}(G).$$

(vi) $\hat{W}_{\alpha \vee \beta}(G) \subseteq \hat{W}_\alpha(G) \cap \hat{W}_\beta(G)$, $\hat{W}_{\Phi \vee \Psi}(G) \subseteq \hat{W}_\Phi(G) \cap \hat{W}_\Psi(G)$.

(vii) $\hat{W}_\alpha(G) \subseteq W_\alpha(G)$.

(viii) *Let G be a group. Then*

$$\hat{W}_\Phi(G) = \prod_{\phi \in \Phi} \hat{W}_\phi(G).$$

Proof. (i) and (ii) follow from the fact that $\mathbf{W}_\alpha$, $\hat{\mathbf{W}}_\alpha$ and $\hat{\mathbf{W}}_\Phi$ are verbal subalgebras of the relatively free algebra $\mathbf{A}$ and $\rho: \mathbf{A} \to G$ is an epimorphism.

(iii) Let x be an arbitrary element of G. Since ρ is epi there exists $\mathbf{x} \in \mathbf{A}$ such that $\mathbf{x}\rho = x$. But then $\sigma(\mathbf{x}) \geqslant 1$, $\hat{\sigma}(\mathbf{x}) \succcurlyeq 1$ and $\Sigma(\mathbf{x}) \succcurlyeq \{1\}$ so x is an element of $W_1(G)$, $\hat{W}_1(G)$ and $\hat{W}_{\{1\}}(G)$.

(iv) follows immediately from the definitions.

(v) Let $x = [a, b]$ where $a \in \hat{W}_\Phi(G)$ and $b \in \hat{W}_\Psi(G)$.

Then there exist $\mathbf{a}$ and $\mathbf{b} \in \mathbf{A}$ such that $\mathbf{a}\rho = a$, $\mathbf{b}\rho = b$, $\Sigma(\mathbf{a}) \succcurlyeq \Phi$ and $\Sigma(\mathbf{b}) \succcurlyeq \Psi$. But then $\Sigma([\mathbf{a}, \mathbf{b}]) \succcurlyeq \Phi + \Psi$ and $x = [\mathbf{a}, \mathbf{b}]\rho$ so $x \in \hat{W}_{\Phi+\Psi}(G)$. Thus the set

$$\{[a, b]: a \in \hat{W}_\Phi(G),\ b \in \hat{W}_\Psi(G)\}$$

is a subset of $\hat{W}_{\Phi+\Psi}(G)$. But $[\hat{W}_\Phi(G), \hat{W}_\Psi(G)]$ is the subalgebra of G generated by this set and is thus contained in $\hat{W}_{\Phi+\Psi}(G)$ also. That

$$[\hat{W}_\alpha(G), \hat{W}_\beta(G)] \subseteq \hat{W}_{\alpha+\beta}(G) \quad \text{and} \quad [W_\alpha(G), W_\beta(G)] \subseteq W_{\alpha+\beta}(G)$$

may now be inferred as described at the end of § 2.

(vi) is a corollary of (iv) using lemma 2·2 (i).

(vii) follows from the fact, already remarked, that for any expression $\mathbf{x} \in \mathbf{A}$, $\hat{\sigma}(\mathbf{x}) \preccurlyeq \sigma(\mathbf{x})$.

(viii) Since $\hat{W}_\phi(G) = \hat{W}_{\{\phi\}}(G)$ by lemma 2·2 (v) and, for any $\phi \in \Phi$, $\{\phi\} \succcurlyeq \Phi$, it follows that $\prod_{\phi \in \Phi} \hat{W}_\phi(G) \subseteq \hat{W}_\Phi(G)$.

To prove the converse inclusion it is sufficient to show that, for $\mathbf{x} \in \mathbf{A}$ such that $\Sigma(\mathbf{x}) \succcurlyeq \Phi$,

$$\mathbf{x}\rho \in \prod_{\phi \in \Phi} \hat{W}_\phi(G)$$

and this follows from the definition of Σ by an easy induction over $ht(\mathbf{x})$.

4. *A lemma on partially ordered groupoids*

Let G be a set, partially ordered by $\preccurlyeq$. Two elements a and b are *comparable* if either $a \preccurlyeq b$ or $b \preccurlyeq a$, otherwise they are *incomparable*. A subset X of G is *totally unordered* if every pair of distinct elements of X are incomparable. The partial order $\preccurlyeq$ is a *partial well-order* if every *non-weakly ascending* sequence, that is every sequence $x_1, x_2, \ldots$ for which $i < j \Rightarrow x_i \not\preccurlyeq x_j$ is finite or, equivalently, if it satisfies the descending chain condition and every totally unordered subset is finite. The equivalence of these definitions follows from the theorem:

Let X be a partially ordered set satisfying the ascending and descending chain conditions and in which every totally unordered subset is finite. Then X is finite.

This theorem is usually attributed to Ramsey and follows from his theorem A (1928) upon setting $r = \mu = 2$ and C_1 and C_2 to be the sets of pairs whose members are comparable and incomparable respectively.

LEMMA 4·1. *Let G be an (additively written) finitely generated groupoid, partially ordered by a relation $\leqslant$ which satisfies*

(i) *the usual regularity condition*

$$a \leqslant b \Rightarrow a+x \leqslant b+x \quad \text{and} \quad x+a \leqslant x+b,$$

(ii) *the additional condition*

$$a \leqslant a+x \quad \text{and} \quad a \leqslant x+a.$$

Then $\leqslant$ is a partial well-order.

Proof. Let H be a groupoid, partially ordered by $\leqslant$, which is the image of G under an order-preserving epimorphism θ. Then the existence of an infinite non-weakly ascending sequence in H implies the existence of a similar one in G: in other words, if G is partially well-ordered by $\leqslant$ so is H. It may now be assumed that G is an absolutely free groupoid, freely generated by a finite set $\{e_1, e_2, \ldots, e_n\}$ and that $\leqslant$ is the coarsest partial order satisfying conditions (i) and (ii) provided that any such order exists. The length $\lambda(a)$ of an element $a \in G$ may be defined by its properties $\lambda(e_i) = 1$ $(1 \leqslant i \leqslant n)$ and $\lambda(a+b) = \lambda(a)+\lambda(b)$. The relation $\leqslant'$ on G defined by $a \leqslant' b$ if $a = b$ or $\lambda(a) < \lambda(b)$ is then a partial order satisfying conditions (i) and (ii) and consequently $\leqslant$, the coarsest partial order satisfying these conditions, exists and then $a \prec b \Rightarrow a \prec' b \Rightarrow \lambda(a) < \lambda(b)$ so $\leqslant$ satisfies the descending chain condition. The proof now proceeds on these assumptions.

Order the ω-sequences (the sequences indexed by the non-negative integers) in G lexicographically: if $(x_i)_{i\in\omega}$ and $(y_i)_{i\in\omega}$ are two such sequences then $(x_i)_{i\in\omega} \prec (y_i)_{i\in\omega}$ if there exists a non-negative integer i_0 such that $x_{i_0} \prec y_{i_0}$ and $i < i_0 \Rightarrow x_i = y_i$.

Suppose there exists an infinite totally unordered subset in G (from this hypothesis a contradiction will be deduced, proving the lemma since it has already been observed that $\leqslant$ satisfies the descending chain condition). Then the set $\mathscr{T}$ of totally unordered ω-sequences (ω-sequences $(t_i)_{i\in\omega}$ such that $i \neq j \Rightarrow t_i$ and t_j are incomparable) is non-empty.

Define an ω-sequence $(c_i)_{i\in\omega}$ recursively as follows: suppose that the c_j $(j < i)$ have been defined. Let T_i be the set
$$T_i = \{t_i\colon (t_j)_{j\in\omega} \in \mathscr{T}, j < i \Rightarrow t_j = c_j\},$$
and let c_i be any minimal element of T_i. Then for any non-negative integer i, there exists $(t_j)_{j\in\omega} \in \mathscr{T}$ such that $j < i \Rightarrow c_j = t_j$. Thus $(c_i)_{i\in\omega}$ is totally unordered and a minimal member of $\mathscr{T}$: it is thus a minimal totally unordered ω-sequence.

Suppose some of the generators $e_1, e_2, \ldots, e_n$ appear as terms of $(c_i)_{i\in\omega}$. Then since this sequence is totally unordered each of these generators can appear at most once, so the new sequence formed from $(c_i)_{i\in\omega}$ by deleting these terms and 'closing the gap' is still an ω-sequence. Further, all the terms in the new sequence must lie in the subgroupoid generated by the remaining generators (those not appearing as terms of $(c_i)_{i\in\omega}$). Thus it may be assumed without loss of generality that none of the generators $e_1, e_2, \ldots, e_n$ appear as terms of $(c_i)_{i\in\omega}$. Then, since G is a free groupoid, each c_i may be written uniquely in the form

$c_i = a_i + b_i$. Let $A = \{a_i\}_{i \in \omega}$ and $B = \{b_i\}_{i \in \omega}$, but notice that the a_i need not be distinct, nor need the b_i be.

Suppose that A contains an infinite totally unordered subset. Then there exists an ω-sequence $(m(i))_{i \in \omega}$ of non-negative integers such that $(a_{m(i)})_{i \in \omega}$ is totally unordered. Now let k be the first non-negative integer for which c_k is comparable with some term of $(a_{m(i)})_{i \in \omega}$: such a k exists since, for instance, $c_{m(0)}$ is comparable with $a_{m(0)}$. Let $a_{m(l)}$ be the term comparable with c_k. If $c_k \leqslant a_{m(l)}$ then $c_k \prec a_{m(l)} + b_{m(l)} = c_{m(l)}$ which is impossible since $(c_i)_{i \in \omega}$ is totally unordered; thus $c_k \succ a_{m(l)}$. Now define a sequence $(d_i)_{i \in \omega}$ by

$$\begin{aligned} d_i &= c_i \quad (i < k) \\ &= a_{m(i+l-k)} \quad (i \geqslant k). \end{aligned}$$

Then $(d_i)_{i \in \omega}$ is totally unordered for if $i < j < k$ then $d_i = c_i$ and $d_j = c_j$ are incomparable since $(c_i)_{i \in \omega}$ is totally unordered, if $i < k \leqslant j$ then $d_i = c_i$ and $d_j = a_{m(j+l+k)}$ are incomparable by the choice of k and if $k \leqslant i < j$ then $d_i = a_{m(i+l-k)}$ and $d_j = a_{m(j+l-k)}$ are incomparable since $(a_{m(i)})_{i \in \omega}$ is totally unordered. However, $i < k \Rightarrow d_i = c_i$ and $d_k = a_{m(l)} \prec c_k$ so $(a_{m(i)})_{i \in \omega} \prec (c_i)_{i \in \omega}$. This contradicts the fact that $(c_i)_{i \in \omega}$ is a minimal totally unordered ω-sequence, thus proving that A contains no infinite totally unordered subsets. By the same argument every totally unordered subset of B is finite. But $\leqslant$ satisfies the descending chain condition so both A and B are partially well-ordered by $\leqslant$.

Now suppose there exists an infinite ascending chain in A. Then there is an ω-sequence $(n(i))_{i \in \omega}$ of non-negative integers such that $i < j \Rightarrow a_{n(i)} \prec a_{n(j)}$. Now consider the sequence $(b_{n(i)})_{i \in \omega}$. Suppose for some $i < j$, $b_{n(i)} \leqslant b_{n(j)}$. Then $c_{n(i)} = a_{n(i)} + b_{n(i)} \prec a_{n(j)} + b_{n(j)} = c_{n(j)}$ which is impossible since $(c_i)_{i \in \omega}$ is totally unordered. Thus, for all $i < j$, $b_{n(i)} \not\leqslant b_{n(j)}$ which is also impossible since B is partially well-ordered and hence contains no infinite non-weakly ascending chains. This proves that A satisfies the ascending chain condition. But it has already been observed that it satisfies the descending chain condition and contains no infinite totally unordered subsets. Thus, by the theorem of Ramsey quoted above, A is finite. By the same argument B is finite also.

But each c_i is the sum of an element of A and one of B, so the c_i are not all distinct. This contradicts the fact that $(c_i)_{i \in \omega}$ is totally unordered and the lemma is proved.

It should be remarked here that although the Axiom of Choice appears to have been used in several places in this proof (most obviously in the definition of $(c_i)_{i \in \omega}$ and less obviously in the assumption of the existence of ω-sequences of distinct elements in infinite subsets of G) this is not in fact the case—for any algebra with a finite or well-ordered set of generators and a finite or well-ordered set of operators can be well-ordered and then a choice function for its non-empty subsets constructed without use of the Axiom of Choice.

Corollary. (i) *The coarse order on a shape range W is a partial well-order and hence a complete lattice order in the sense that, for any subset Φ of W the meet $\wedge \Phi$ and join $\vee \Phi$ exist.* (ii) *Any totally unordered subset Φ of W is finite and the relation $\leqslant$ between totally unordered subsets of W (definition* 2·2) *satisfies the descending chain condition.*

5. *Shape subgroups as products of commutator subgroups*

For any group G the subgroups formed from G by repeated subgroup commutation, together with the trivial subgroup $\{1\}$, will be called the *commutator subgroups* of G; in other words, the set of commutator subgroups of G is the smallest set of subgroups of G which has G itself and $\{1\}$ as members and is closed under the operation: A and B are commutator subgroups $\Rightarrow [A, B]$ is also. The object of this section is to show how, for a given shape range W, the corresponding shape subgroups can be described in ordinary group theoretic terms as finite products of commutator subgroups in G.

Definition 5·1. *For any element ϕ of a shape range W, $U(\phi)$ is the set of all pairs (α, β) for which*

(i) $\alpha, \beta \in W$,

(ii) $\alpha+\beta \geqslant \phi$,

(iii) $\beta \leqslant \alpha < \phi$,

(iv) $\alpha' < \alpha \Rightarrow \alpha'+\beta < \phi$ *and* $\beta' < \beta \Rightarrow \alpha+\beta' < \phi$,

and $\hat{U}(\phi)$ *is the set of all pairs* (α, β) *for which*

(i)′ $\alpha, \beta \in W$,

(ii)′ $\alpha+\beta \geqslant \phi$,

(iii)′ $\beta \leqslant \alpha \ngeqslant \phi$,

(iv)′ $\alpha' \prec \alpha \Rightarrow \alpha'+\beta \ngeqslant \phi$ *and* $\beta' \prec \beta \Rightarrow \alpha+\beta' \ngeqslant \phi$.

Notice that $U(1)$, $\hat{U}(1)$, $U(\infty)$ and $\hat{U}(\infty)$ are all empty.

Lemma 5·1. *Let ϕ be an element of a shape range W other than* 1 *and let G be a group. Then*

$$\hat{W}_\phi(G) = \prod_{(\alpha,\beta)\in\hat{U}(\phi)} [\hat{W}_\alpha(G), \hat{W}_\beta(G)].$$

Proof. If $\phi = \infty$ the result is trivial, using the usual convention that the product of an empty set of subgroups of G is the trivial subgroup. Now suppose $\phi \neq 1$ or ∞.

Suppose $(\alpha, \beta) \in \hat{U}(\phi)$. Then

$$\begin{aligned}[\hat{W}_\alpha(G), \hat{W}_\beta(G)] &\leqslant \hat{W}_{\alpha+\beta}(G) \quad \text{by theorem 3·1 (v)},\\ &\leqslant \hat{W}_\phi(G) \quad \text{by theorem 3·1 (iv)},\end{aligned}$$

since $\alpha+\beta \geqslant \phi$. Thus
$$\prod_{(\alpha,\beta)\in\hat{U}(\phi)} [\hat{W}_\alpha(G), \hat{W}_\beta(G)] \leqslant \hat{W}_\phi(G).$$

To prove the converse inclusion, notice first that if $\alpha', \beta' \in W$, $\alpha' \ngeqslant \phi$, $\beta' \ngeqslant \phi$ and $\alpha'+\beta' \geqslant \phi$ then there exist $\alpha, \beta \in W$ such that $\alpha \preccurlyeq \alpha'$, $\beta \preccurlyeq \beta'$ and either (α, β) or $(\beta, \alpha) \in \hat{U}(\phi)$: this follows immediately from definition 5·1 by induction over ϕ.

Let ρ: $\mathbf{A} \to G$ be a description of G and $\hat{\sigma}$: $\mathbf{A} \to W$ the appropriate coarse shape. It is now sufficient to show that, for $\mathbf{x} \in \mathbf{A}$,

$$\hat{\sigma}(\mathbf{x}) \geqslant \phi \Rightarrow \mathbf{x}\rho \in \prod_{(\alpha,\beta)\in\hat{U}(\phi)} [\hat{W}_\alpha(G), \hat{W}_\beta(G)],$$

by induction over the height of $\mathbf{x}$. If $ht(\mathbf{x}) = 1$ then either $\mathbf{x} = \mathbf{1}$, in which case the result is trivially true, or else $\mathbf{x} = \mathbf{g}_i \in \mathbf{G}$, in which case $\phi \preccurlyeq \hat{\sigma}(\mathbf{x}) = 1$ contrary to the assumption. Now suppose $ht(\mathbf{x}) > 1$ and the result is true for all $\mathbf{x}$ of smaller height. Then there are three possibilities: if $\mathbf{x} = \mathbf{u}^{-1}$ or $\mathbf{x} = \mathbf{x}_1\mathbf{x}_2$ the result is true since

$$\prod_{(\alpha,\beta)\in\hat{U}(\phi)} [\hat{W}_\alpha(G), \hat{W}_\beta(G)]$$

is a subgroup and if $\mathbf{x} = [\mathbf{a}, \mathbf{b}]$ then $\alpha'+\beta' = \hat{\sigma}(\mathbf{x}) \geqslant \phi$ where $\alpha' = \sigma(\mathbf{a})$ and $\beta' = \hat{\sigma}(\mathbf{b})$. Now if $\alpha' \geqslant \phi$ then by the inductive hypothesis

$$\mathbf{a}\rho \in \prod_{(\alpha,\beta)\in\hat{U}(\phi)} [\hat{W}_\alpha(G), \hat{W}_\beta(G)]$$

and then, since this subgroup is normal, so does $\mathbf{x}\rho = [\mathbf{a}\rho, \mathbf{b}\rho]$. If $\beta' \geqslant \phi$ the same argument applies. Finally, if $\alpha' \ngeqslant \phi$ and $\beta' \ngeqslant \phi$, then, as observed above, there exist $\alpha, \beta \in W$ such that $\alpha \leqslant \alpha'$, $\beta \leqslant \beta'$ and either (α, β) or $(\beta, \alpha) \in \hat{U}(\phi)$. Then $\mathbf{a}\rho \in \hat{W}_\alpha(G)$ and $\mathbf{b}\rho \in \hat{W}_\beta(G)$ so

$$\mathbf{x}\rho = [\mathbf{a}\rho, \mathbf{b}\rho] \in [\hat{W}_\alpha(G), \hat{W}_\beta(G)] = [\hat{W}_\beta(G), \hat{W}_\alpha(G)]$$

and the result is true.

Translating this lemma according to the metatheorem of § 2 yields:

Corollary 1. *Let ϕ be an element of a shape range W other than 1 and let G be a group. Then*

$$W_\phi(G) = \prod_{(\alpha,\beta) \in U(\phi)} [W_\alpha(G), W_\beta(G)].$$

Corollary 2. *Let ϕ be an element of a shape range W and G be a group. Then $W_\phi(G)$ and $\hat{W}_\phi(G)$ are each a product of a finite set of commutator subgroups of G.*

Proof. That $\hat{W}_\phi(G)$ is the product of some set of commutator subgroups of G follows from the statement of the lemma by induction over ϕ. But every product of commutator subgroups of G is the product of a finite number of them because the set of commutator subgroups is a groupoid under subgroup-commutation, finitely generated (by G and $\{1\}$) and ordered by inclusion which satisfies $A \supseteq [A, B]$ and $B \supseteq [A, B]$, so by Lemma 4·1, every collection of commutator subgroups in which no one is a proper subgroup of another is necessarily finite. Translation of this result by the metatheorem yields the corresponding one for $W_\phi(G)$.

Returning to the lemma and corollary 1, the formulae given there provide a practical recursive method for calculating ordinary group-theoretic expressions for the shape subgroups corresponding to the shape range W once the structure of W is known. This process is applied in the next lemma to the shape range N^- and in chapter III to more interesting ones.

Lemma 5·2. *For each positive integer c and any group G,*

$$N_c^-(G) = \hat{N}_c^-(G) = \lambda_c(G).$$

Proof. First, since the fine and coarse orders on N^- are the same, it follows that

$$N_c^-(G) = \hat{N}_c^-(G).$$

The proof that $N_c^-(G) = \gamma_c(G)$ is by induction over c. When $c = 1$ this is trivial. For $c > 1$ it follows by definitions 1·3 and 5·1 that $U(c) = \{(c-r, r) : 1 \leqslant r \leqslant \frac{1}{2}c\}$ so

$$\begin{aligned} N_c^-(G) &= \prod_{1 \leqslant r \leqslant \frac{1}{2}c} [N_{c-r}^-(G), N_r^-(G)] \\ &= \prod_{1 \leqslant r \leqslant \frac{1}{2}c} [\gamma_{c-r}(G), \gamma_r(G)] \\ &= \gamma_c(G). \end{aligned}$$

Corollary. *For each positive integer c,*

$$\mathfrak{N}_c^- = \hat{\mathfrak{N}}_c^- = \mathfrak{N}_{c-1},$$

the variety of all groups which are nilpotent of class $c-1$.

These calculations show that routine application of lemma 5·1 may yield a redundant product and corollary 2 provides no guarantee that this may not be infinite (as written);

but for applications this could well be inconvenient. Thus the following lemma is of practical interest, although it is not used in the development of the theory.

LEMMA 5·3. *For each element ϕ of a shape range W the sets $\hat{U}(\phi)$ and $U(\phi)$ are finite.*

Proof. It will be proved that $\hat{U}(\phi)$ is finite: that $U(\phi)$ is finite then follows as usual by application of the metatheorem of § 2. Write $A = \{\alpha\colon (\alpha,\beta) \in \hat{U}(\phi)\}$ and $B = \{\beta\colon (\alpha,\beta) \in \hat{U}(\phi)\}$. By the corollary to lemma 4·1, both A and B are partially well-ordered by $\leqslant$ so in order to show that A is finite it is sufficient to show that it satisfies the ascending chain condition. Suppose then that A contains an infinite ascending chain. Then there is an infinite sequence $(\alpha_1,\beta_1), (\alpha_2,\beta_2) \ldots$ in $U(\phi)$ such that $\alpha_1 \prec \alpha_2 \prec \ldots$. Now suppose, for some $i<j$, $\beta_i \leqslant \beta_j$. Then $\alpha_i+\beta_j \geqslant \alpha_i+\beta_i \geqslant \phi$ and $\alpha_i \prec \alpha_j$ which contradicts $(\alpha_j,\beta_j) \in \hat{U}(\phi)$. Thus the sequence $\beta_1, \beta_2 \ldots$ is an infinite non-weakly ascending sequence which is impossible since $\leqslant$ is a partial well-order. Thus A is finite; by the same argument B is finite and then so is $\hat{U}(\phi)$.

6. *Basic commutators*

In this section a well-ordering of the set **C** of commutators is described. This order depends upon two things: the order imposed upon the set **G** of generators by the ordinals indexing it and the shape $\sigma\colon \mathbf{A} \to W$. Subsequently a subset of **C**, the set of *W-basic commutators*, is defined in terms of this. The group-theoretic results arrived at later will not depend upon the well-ordering of **G**, though they do depend very much on W; however, the intermediate steps rely heavily on this well-ordering.

For the remainder of this chapter the word 'shape' will refer to 'fine shape'.

DEFINITION 6·1. *Let $\sigma\colon \mathbf{A} \to W$ be a shape. The relation $<$ on* **C** *is defined*:

(i) *If $\mathfrak{g}_i$ and $\mathfrak{g}_j$ are generators (commutators of shape 1) then*

$$\mathfrak{g}_i < \mathfrak{g}_j \Leftrightarrow i < j.$$

(ii) *If* **a** *and* **b** *are commutators of distinct shapes ($\sigma(\mathbf{a}) \neq \sigma(\mathbf{b})$) then*

$$\mathbf{a} < \mathbf{b} \Leftrightarrow \sigma(\mathbf{a}) < \sigma(\mathbf{b}).$$

(iii) *If* **a** *and* **b** *are commutators of the same shape, $\sigma(\mathbf{a}) = \sigma(\mathbf{b}) = \xi$ say, and $\xi \neq 1$ then '$\mathbf{a} < \mathbf{b}$' is defined recursively over ξ. An intermediate definition must be made: suppose* **x** *is any commutator of shape ξ. Then, since $\xi > 1$, it may be written in the form $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2]$. Then the* leading *and* trailing parts *of* **x** *are*

$$\begin{aligned} \mathrm{ld}\,(\mathbf{x}) &= \mathbf{x}_1 \quad \textit{if} \quad \mathbf{x}_2 < \mathbf{x}_1 \quad \textit{or} \quad \mathbf{x}_2 = \mathbf{x}_1, \\ &= \mathbf{x}_2 \quad \textit{otherwise; and} \\ \mathrm{tr}\,(\mathbf{x}) &= \mathbf{x}_2 \quad \textit{if} \quad \mathbf{x}_2 < \mathbf{x}_1 \quad \textit{or} \quad \mathbf{x}_2 = \mathbf{x}_1, \\ &= \mathbf{x}_1 \quad \textit{otherwise}. \end{aligned}$$

Then $\mathbf{a} < \mathbf{b}$ *if and only if*

(*a*) $\mathrm{ld}\,(\mathbf{a}) < \mathrm{ld}\,(\mathbf{b})$,

(*b*) $\mathrm{ld}\,(\mathbf{a}) = \mathrm{ld}\,(\mathbf{b})$ *and* $\mathrm{tr}\,(\mathbf{a}) < \mathrm{tr}\,(\mathbf{b})$,

or (*c*) $\mathrm{ld}\,(\mathbf{a}) = \mathrm{ld}\,(\mathbf{b})$, $\mathrm{tr}\,(\mathbf{a}) = \mathrm{tr}\,(\mathbf{b})$ *and* $\mathbf{b}_1 < \mathbf{a}_1$

(where $\mathbf{a} = [\mathbf{a}_1, \mathbf{a}_2]$ and $\mathbf{b} = [\mathbf{b}_1, \mathbf{b}_2]$). The reversal of the relation in part (iii *c*) *is intentional. With regard to this part of the definition it will be observed that if* $\mathrm{ld}\,(\mathbf{a}) = \mathrm{ld}\,(\mathbf{b})$ *and* $\mathrm{tr}\,(\mathbf{a}) = \mathrm{tr}\,(\mathbf{b})$ *then either* $\mathbf{b} = \mathbf{a}$ *or* $\mathbf{b} = [\mathbf{a}_2, \mathbf{a}_1]$.

(iv) *The relation* $\leqslant$ *on* $\mathbf{C}$ *is defined*: $\mathbf{a} \leqslant \mathbf{b}$ *if* $\mathbf{a} < \mathbf{b}$ *or* $\mathbf{a} = \mathbf{b}$. *Subject to the proof in the next lemma that* $\leqslant$ *is an order for* $\mathbf{C}$, *the notations* $\geqslant$ *and* $>$ *will be used in the usual sense and the relation* $\leqslant$ *called the* W-*ordering of* $\mathbf{C}$.

LEMMA 6·1. *With the notation of definition* 6·1, $\leqslant$ *is a (full) well-ordering of* $\mathbf{C}$.

Proof. Notice first that

(α) $\mathbf{a} \leqslant \mathbf{b} \Rightarrow \sigma(\mathbf{a}) \leqslant \sigma(\mathbf{b})$,

(β) $\mathfrak{g}_i \leqslant \mathfrak{g}_j \Rightarrow i \leqslant j$,

(γ) $\mathbf{a} \leqslant \mathbf{b}$ and $\sigma(\mathbf{a}) = \sigma(\mathbf{b}) > 1 \Rightarrow \mathrm{ld}\,(\mathbf{a}) \leqslant \mathrm{ld}\,(\mathbf{b})$,

(δ) $\mathbf{a} \leqslant \mathbf{b}$, $\sigma(\mathbf{a}) = \sigma(\mathbf{b}) > 1$ and $\mathrm{ld}\,(\mathbf{a}) = \mathrm{ld}\,(\mathbf{b}) \Rightarrow \mathrm{tr}\,(\mathbf{a}) \leqslant \mathrm{tr}\,(\mathbf{b})$,

(ϵ) $\mathbf{a} \leqslant \mathbf{b}$, $\sigma(\mathbf{a}) = \sigma(\mathbf{b}) > 1$, $\mathrm{ld}\,(\mathbf{a}) = \mathrm{ld}\,(\mathbf{b})$ and $\mathrm{tr}\,(\mathbf{a}) = \mathrm{tr}\,(\mathbf{b}) \Rightarrow \mathbf{b}_1 \leqslant \mathbf{a}_1$ (where $\mathbf{a} = [\mathbf{a}_1, \mathbf{a}_2]$ and $\mathbf{b} = [\mathbf{b}_1, \mathbf{b}_2]$).

Since it is not yet proved that $\leqslant$ is a partial order, these statements must be proved by checking the various possibilities listed in the definition. It is now shown that $\leqslant$ is indeed a partial order.

(i) $\leqslant$ is reflexive, by part (iv) of the definition.

(ii) $\leqslant$ is weakly antisymmetric. Suppose that $\mathbf{x} \leqslant \mathbf{y}$ and $\mathbf{y} \leqslant \mathbf{x}$. Then, by ($\alpha$), $\sigma(\mathbf{x}) = \sigma(\mathbf{y})$ since the relation $\leqslant$ on W is known to be an order. Write $\sigma(\mathbf{x}) = \sigma(\mathbf{y}) = \xi$. If $\xi = 1$ then there exist ordinals i and j such that $\mathbf{x} = \mathfrak{g}_i$ and $\mathbf{y} = \mathfrak{g}_j$ and then $i = j$ by (β) so $\mathbf{x} = \mathbf{y}$. If $\xi > 1$ it may be supposed inductively that the relation $\leqslant$ is weakly antisymmetric on the set of all commutators of shape $< \xi$. Then, by (γ), $\mathrm{ld}\,(\mathbf{x}) = \mathrm{ld}\,(\mathbf{y})$ so, by (δ), $\mathrm{tr}\,(\mathbf{x}) = \mathrm{tr}\,(\mathbf{y})$ and finally by (ϵ), $\mathbf{x}_1 = \mathbf{y}_1$ where $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2]$ and $\mathbf{y} = [\mathbf{y}_1 \, \mathbf{y}_2]$. Hence $\mathbf{x} = \mathbf{y}$.

(iii) $\leqslant$ is transitive. Suppose $\mathbf{x} \leqslant \mathbf{y}$ and $\mathbf{y} \leqslant \mathbf{z}$. Then by ($\alpha$), $\sigma(\mathbf{x}) \leqslant \sigma(\mathbf{y}) \leqslant \sigma(\mathbf{z})$. If $\sigma(\mathbf{x}) < \sigma(\mathbf{z})$ then $\mathbf{x} < \mathbf{z}$ by part (ii) of the definition. Otherwise $\sigma(\mathbf{x}) = \sigma(\mathbf{y}) = \sigma(\mathbf{z}) = \xi$ say. From now on the proof follows the same pattern as that of weak antisymmetry, helped by the fact that it may be assumed inductively that the relation $\leqslant$ is a partial order on the set of all commutators of shape $< \xi$.

(iv) $\leqslant$ is a (full) well-ordering of $\mathbf{C}$. Suppose $\mathbf{X}$ is a non-empty subset of $\mathbf{C}$. It is shown that $\mathbf{X}$ has a least element. Let $\mathbf{X}_1$ be the set of commutators of least shape, ξ say, in $\mathbf{X}$. This is non-empty and, by part (ii) of the definition, if $\mathbf{X}_1$ has a least element so does X. If $\xi = 1$ then $\mathbf{X}_1$ is a non-empty subset of $\mathbf{G}$ which is well-ordered by part (i) of the definition, so $\mathbf{X}_1$ has a least element. If $\xi > 1$ it may be assumed inductively that the set of all commutators of shape $< \xi$ is well-ordered by $\leqslant$. Then the set $\mathbf{X}_2$ of all commutators of least leading part, $\mathbf{l}$ say, in $\mathbf{X}_1$ is non-empty and if it has a least element so does $\mathbf{X}$. Further, the set $\mathbf{X}_3$ of all commutators of least trailing part, $\mathbf{t}$ say, in $\mathbf{X}_2$ is non-empty and if it has a least element so does $\mathbf{X}$. But $\mathbf{X}_3 \subseteq \{[\mathbf{l}, \mathbf{t}], [\mathbf{t}, \mathbf{l}]\}$ and so has a least element since $[\mathbf{l}, \mathbf{t}] \leqslant [\mathbf{t}, \mathbf{l}]$.

DEFINITION 6·2. (A) *Let* $\sigma\colon \mathbf{A} \to W$ *be a shape. A particular type of commutator, called a* W-basic commutator *is defined recursively over its weight by*

(i) *Every commutator of weight* 1 (*that is, every member of* $\mathbf{G}$) *is a* W-*basic commutator.*

(ii) *A commutator* $\mathbf{c} = [\mathbf{b}, \mathbf{a}]$ *of weight* > 1 *is* W-*basic if* (*a*) $\mathbf{a}$ *and* $\mathbf{b}$ *are both* W-*basic commutators*, (*b*) $\mathbf{a} < \mathbf{b}$ (*under the* W-*ordering of* $\mathbf{C}$), *and* (*c*) *if* $\mathbf{b} = [\mathbf{b}_1, \mathbf{b}_2]$ *then* $\mathbf{b}_2 \leqslant \mathbf{a}$.

(B) *A* W-*basic expression is an expression of the form* $\mathbf{1}$ *or* $\mathbf{b}_1^{\alpha_1} \mathbf{b}_2^{\alpha_2} \ldots \mathbf{b}_k^{\alpha_k}$ *where*

(i) k *is a positive integer*,

(ii) *each* $\mathbf{b}_i$ *is a* W-*basic commutator*,

(iii) $\mathbf{b}_1 < \mathbf{b}_2 < \ldots < \mathbf{b}_k$ *under the W-ordering, and*

(iv) *each* α_i *is a non-zero integer (positive or negative).*

The set of all W-basic expressions is denoted $\mathbf{B}^W$.

Whenever there is no doubt about which shape range W is under consideration, the adjective 'W-basic' will be abbreviated to 'basic' and the set $\mathbf{B}^W$ *written simply* $\mathbf{B}$.

For any $\alpha \in W$, *the set of basic expressions in which*

(v) $\sigma(\mathbf{b}_i) < \alpha (1 \leqslant i \leqslant k)$ *is denoted* $\mathbf{B}^W_\alpha$ *or simply* $\mathbf{B}_\alpha$ *and the set of basic expressions in which*

(v)′ $\hat{\sigma}(\mathbf{b}_i) \ngeqslant \alpha (1 \leqslant i \leqslant k)$ *is denoted* $\hat{\mathbf{B}}^W_\alpha$ *or simply* $\hat{\mathbf{B}}_\alpha$.

For any subset Φ *of W the set of basic expressions in which*

(v)″ $\phi \in \Phi \Rightarrow \hat{\sigma}(\mathbf{b}_i) \ngeqslant \phi (1 \leqslant i \leqslant k)$ *is denoted* $\hat{\mathbf{B}}^W_\Phi$ *or simply* $\hat{\mathbf{B}}_\Phi$.

For any positive integer c the set of basic expressions in which

(v)‴ $\mathrm{wt}\,(\mathbf{b}_i) \leqslant c\ (1 \leqslant i \leqslant k)$ *is denoted* $\mathbf{B}^W_{(c)}$ *or simply* $\mathbf{B}_{(c)}$.

The expression $\mathbf{1}$ is assumed to be a member of all these sets.

Notice that, if $\mathbf{u}$ and $\mathbf{v}$ are basic expressions other than $\mathbf{1}$, $\mathbf{u} \in \mathbf{B}_\alpha$ and $\sigma(\mathbf{v}) \geqslant \alpha$ then $\mathbf{uv}$ is also a basic expression.

7. *The number of basic commutators*

This section is devoted to finding an expression for the number of basic commutators of a given weight when the number τ of generators is finite. The argument given here is a modified form of Witt's original one (1937); his argument does not carry over exactly since it requires the order type of the set of basic commutators to be ω.

Definition 7·1. (A). *An order* $\leqslant$ *defined on the set* $\mathbf{C}$ *of commutators is a B-*order *if*

(i) $\leqslant$ *(fully) well-orders* $\mathbf{C}$,

(ii) $\mathbf{a} < [\mathbf{a}, \mathbf{b}]$ *and* $\mathbf{b} < [\mathbf{a}, \mathbf{b}]$,

(iii) $\mathbf{g}_i < [\mathbf{a}, \mathbf{b}]$ *and* $i < j \Rightarrow \mathbf{g}_i < \mathbf{g}_j$.

(B). *If* $\leqslant$ *is any B-*order, *then a* $(\leqslant)$-basic commutator *is defined recursively over its weight:*

(i) *Every* $\mathbf{g}_i \in \mathbf{G}$ *is a* $(\leqslant)$*-basic commutator,*

(ii) $[\mathbf{b}, \mathbf{a}]$ *is a* $(\leqslant)$*-basic commutator if (a) both* $\mathbf{a}$ *and* $\mathbf{b}$ *are* $(\leqslant)$*-basic commutators, (b)* $\mathbf{a} < \mathbf{b}$ *and (c) if* $\mathbf{b} = [\mathbf{b}_1, \mathbf{b}_2]$ *then* $\mathbf{b}_2 \leqslant \mathbf{a}$.

Until further notice it will be assumed that there is a fixed B-order $\leqslant$ defined on the set $\mathbf{C}$ of commutators, in terms of which the following definitions are made.

Definition 7·2.

(ii) *A commutator* $\mathbf{c}$ *is* $\mathbf{b}$-compatible, *where* $\mathbf{b}$ *is any* $(\leqslant)$*-basic commutator, if* $[\mathbf{c}, \mathbf{b}]$ *is a* $(\leqslant)$*-basic commutator.*

(ii) *For any* $(\leqslant)$*-basic commutator* $\mathbf{b}$, *write* $\mathbf{b}^+$ *for the successor of* $\mathbf{b}$ *under the restriction of* $\leqslant$ *to the set of* $(\leqslant)$*-basic commutators.*

Definition 7·3. *For any expressions* $\mathbf{b}, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ *and non-negative integers* $n_1, n_2, \ldots, n_k$ *the expression* $[\mathbf{b}, n_1\mathbf{a}_1, n_2\mathbf{a}_2, \ldots, n_k\mathbf{a}_k]$ *is defined recursively by*

$$[\mathbf{b}, 0\mathbf{a}] = \mathbf{b},$$

$$\begin{aligned}[\mathbf{b}, n_1\mathbf{a}_1, n_2\mathbf{a}_2, \ldots, n_k\mathbf{a}_k] &= [\mathbf{b}, n_1\mathbf{a}_1, n_2\mathbf{a}_2, \ldots, n_{k-1}\mathbf{a}_{k-1}] \quad \text{if} \quad n_k = 0,\\ &= [[\mathbf{b}, n_1\mathbf{a}_1, n_2\mathbf{a}_2, \ldots, (n_k - 1)\mathbf{a}_k], \mathbf{a}_k] \quad \text{if} \quad n_k > 0.\end{aligned}$$

If any of the n_i *are* 1 *they may be omitted.*

It follows immediately from these definitions that

LEMMA 7·1.

(i) *If* $\mathbf{a}$ *and* $\mathbf{b}$ *are commutators, then* $\mathrm{wt}\,([\mathbf{b}, n\mathbf{a}]) = \mathrm{wt}\,(\mathbf{b}) + n\,\mathrm{wt}\,(\mathbf{a})$,

(ii) *if* $\mathbf{b}$ *is* $(\leqslant)$*-basic and* $\mathbf{c}$ *is* $\mathbf{b}$*-compatible, then* $[\mathbf{c}, n\mathbf{b}]$ *is* $(\leqslant)$*-basic and* $\mathbf{b}$*-compatible for all* $n \geqslant 0$ *and* $\mathbf{b}^+$*-compatible for all* $n \geqslant 1$.

LEMMA 7·2. *If the order-type of the set* $\mathbf{C}$ *of commutators under* $\leqslant$ *is* ω *and* τ *is finite then the collection of* $(\leqslant)$*-basic commutators may be indexed by the positive integers* $\{\mathbf{b}_i\}_{i=1}^{\infty}$ *so that* $i < j \Leftrightarrow \mathbf{b}_i < \mathbf{b}_j$ *and then, for each positive integer* w*, there exists an integer* n_w *such that* $i \geqslant n_w \Rightarrow \mathrm{wt}\,(\mathbf{b}_i) \geqslant w$.

Proof. Since the number of generators of $\mathbf{A}$ is finite, an easy inductive argument shows that the number of commutators of weight less than any fixed integer w is finite. The lemma follows immediately.

It should be remarked that, as the indices have been defined, $\mathbf{b}_1 = \mathbf{g}_0$, $\mathbf{b}_2 = \mathbf{g}_1$ and so on.

DEFINITION 7·4. *With the conditions of lemma* 7·2*, a sequence* $(\mathbf{X}_i)_{i=0}^{\infty}$ *of subsets of* $\mathbf{A}$*, is defined as follows*:

(i) $\mathbf{X}_0 = \mathbf{G}$, *and*

(ii) *for* $i > 0$, $\mathbf{X}_i$ *is the set of all* $\mathbf{b}_i$*-compatible commutators.*

LEMMA 7·3. *For each integer* $r \geqslant 1$,

(i) $\mathbf{X}_{r-1} - \mathbf{X}_r = \{\mathbf{b}_r\}$; *and*

(ii) $\mathbf{X}_r - \mathbf{X}_{r-1} = \{[\mathbf{c}, n\mathbf{b}_r] : \mathbf{c} \in \mathbf{X}_r \cap \mathbf{X}_{r-1}$ *and* $n \geqslant 1\}$.

Proof. The argument is slightly different for the cases $r = 1$ and $r > 1$.

$r = 1$

(i) $\mathbf{X}_0 = \mathbf{G}$, $\mathbf{b}_1 = \mathbf{g}_0$ and $\mathbf{X}_1$ is the set of all $\mathbf{g}_0$-compatible commutators. Clearly $\mathbf{g}_0 \in \mathbf{G}$ and is not $\mathbf{g}_0$-compatible so $\{\mathbf{g}_0\} \subseteq \mathbf{X}_0 - \mathbf{X}_1$. Now suppose $\mathbf{c} \in \mathbf{X}_0 - \mathbf{X}_1$. Then $\mathbf{c} \in \mathbf{X}_0 = \mathbf{G}$ so there exists $i < \tau$ such that $\mathbf{c} = \mathbf{g}_i$. If $i \geqslant 1$ then $[\mathbf{c}, \mathbf{g}_0] = [\mathbf{g}_i, \mathbf{g}_0]$ is $(\leqslant)$-basic and so $\mathbf{g}_i \in \mathbf{X}_1$ contradicting the choice of $\mathbf{c} = \mathbf{g}_i$. Thus $\mathbf{c} = \mathbf{g}_0$ and $\mathbf{X}_0 - \mathbf{X}_1 = \{\mathbf{g}_0\} = \{\mathbf{b}_1\}$.

(ii) By lemma 7·1 (ii), all commutators of the form $[\mathbf{g}_i, n\mathbf{g}_0]$, where $i \geqslant 1$ and $n \geqslant 0$, are $\mathbf{g}_0$-compatible and thus members of $\mathbf{X}_1$. It is now shown that all members of $\mathbf{X}_1$ are of this form by induction over their weight. If $\mathbf{g}_i \in \mathbf{X}_1$ then it is $\mathbf{g}_0$-compatible so $i \geqslant 1$ and since $\mathbf{g}_i = [\mathbf{g}_i, 0\mathbf{g}_0]$ it is of the desired form. If $[\mathbf{b}, \mathbf{a}] \in \mathbf{X}_1$ then it is $\mathbf{g}_0$-compatible so $[\mathbf{b}, \mathbf{a}, \mathbf{g}_0]$ is $(\leqslant)$-basic and thus $\mathbf{a} \leqslant \mathbf{g}_0$, that is, $\mathbf{a} = \mathbf{g}_0$. Then $[\mathbf{b}, \mathbf{a}] = [\mathbf{b}, \mathbf{g}_0]$ is $(\leqslant)$-basic so $\mathbf{b}$ is $\mathbf{g}_0$-compatible. Inductively, $\mathbf{b} = [\mathbf{g}_i, n\mathbf{g}_0]$ where $i \geqslant 1$ and $n \geqslant 0$ and then

$$[\mathbf{b}, \mathbf{a}] = [\mathbf{g}_i, (n+1)\,\mathbf{g}_0]$$

which is of the desired form. This proves that $\mathbf{X}_1 = \{[\mathbf{g}_1, n\mathbf{g}_0] : i \geqslant 1, n \geqslant 0\}$ and then, since $\mathbf{X}_0 = \mathbf{G}$, $\mathbf{X}_1 - \mathbf{X}_0 = \{[\mathbf{g}_1, n\mathbf{g}_0] : i \geqslant 1, n \geqslant 1\}$ and $\mathbf{X}_1 \cap \mathbf{X}_0 = \{\mathbf{g}_i : i \geqslant 1\}$. This case is proved.

$r > 1$

(i) $\mathbf{b}_r > \mathbf{b}_{r-1}$ and, if $\mathbf{b}_r = [\mathbf{c}_1, \mathbf{c}_2]$ then $\mathbf{c}_2 < \mathbf{b}_r$ so that $\mathbf{c}_2 \leqslant \mathbf{b}_{r-1}$. Thus $[\mathbf{b}_r, \mathbf{b}_{r-1}]$ is $(\leqslant)$-basic so $\mathbf{b}_r \in \mathbf{X}_{r-1}$. But $\mathbf{b}_r$ is not $\mathbf{b}_r$-compatible so $\mathbf{b}_r \in \mathbf{X}_{r-1} - \mathbf{X}_r$. Conversely, suppose that $\mathbf{c} \in \mathbf{X}_{r-1} - \mathbf{X}_r$. Then $[\mathbf{c}, \mathbf{b}_{r-1}]$ is $(\leqslant)$-basic but $[\mathbf{c}, \mathbf{b}_r]$ is not. Thus one of the conditions of definition 7·1 (B) must fail for $[\mathbf{c}, \mathbf{b}_r]$. Now $\mathbf{c}$ is $(\leqslant)$-basic since $[\mathbf{c}, \mathbf{b}_{r-1}]$ is and if $\mathbf{c} = [\mathbf{c}_1, \mathbf{c}_2]$

then $\mathbf{c}_2 \leqslant \mathbf{b}_{r-1}$ for the same reason so $\mathbf{c}_2 < \mathbf{b}_r$. Thus the only condition that can fail is (b), that is, $\mathbf{c} \leqslant \mathbf{b}_r$. But again, since $[\mathbf{c}, \mathbf{b}_{r-1}]$ is ($\leqslant$)-basic, $\mathbf{c} > \mathbf{b}_{r-1}$. Thus $\mathbf{c} = \mathbf{b}_r$. This proves that $\mathbf{X}_{r-1} - \mathbf{X}_r = \{\mathbf{b}_r\}$.

(ii) If $\mathbf{c} \in \mathbf{X}_{r-1} \cap \mathbf{X}_r$ and $n \geqslant 1$ then $[\mathbf{c}, n\mathbf{b}_r] \in \mathbf{X}_r$ by Lemma 7·1 (ii). But, since $n \geqslant 1$, $[\mathbf{c}, n\mathbf{b}_r]$ is not $\mathbf{b}_{r-1}$-compatible. Thus $[\mathbf{c}, n\mathbf{b}_r] \in \mathbf{X}_r - \mathbf{X}_{r-1}$. Conversely, suppose $\mathbf{c} \in \mathbf{X}_r - \mathbf{X}_{r-1}$. The argument proceeds by induction over the weight of $\mathbf{c}$. Now $[\mathbf{c}, \mathbf{b}_r]$ is ($\leqslant$)-basic but $[\mathbf{c}, \mathbf{b}_{r-1}]$ is not, so one of the conditions of definition 7·1 (B) must fail for $[\mathbf{c}, \mathbf{b}_{r-1}]$. But $\mathbf{c}$ is ($\leqslant$)-basic since $[\mathbf{c}, \mathbf{b}_r]$ is and for the same reason $\mathbf{c} > \mathbf{b}_r > \mathbf{b}_{r-1}$. Thus the only condition that can fail is (c), that is, $\mathbf{c} = [\mathbf{c}_1, \mathbf{c}_2]$ and $\mathbf{c}_2 > \mathbf{b}_{r-1}$. But again, since $[\mathbf{c}, \mathbf{b}_r]$ is ($\leqslant$)-basic, $\mathbf{c}_2 \leqslant \mathbf{b}_r$. Thus $\mathbf{c}_2 = \mathbf{b}_r$ and $\mathbf{c} = [\mathbf{c}_1, \mathbf{b}_r]$ so $\mathbf{c}_1$ is $\mathbf{b}_r$-compatible. If $\mathbf{c}_1$ is also $\mathbf{b}_{r-1}$-compatible, then $\mathbf{c} = [\mathbf{c}_1, 1\mathbf{b}_r]$ and $\mathbf{c}_1 \in \mathbf{X}_r \cap \mathbf{X}_{r-1}$ and the result is true. If $\mathbf{c}_1$ is not $\mathbf{b}_{r-1}$-compatible then $\mathbf{c}_1 \in \mathbf{X}_r - \mathbf{X}_{r-1}$ and inductively $\mathbf{c}_1 = [\mathbf{c}', n\mathbf{b}_r]$ where $\mathbf{c}' \in \mathbf{X}_r \cap \mathbf{X}_{r-1}$ and $n \geqslant 1$. Then

$$\mathbf{c} = [\mathbf{c}_1, \mathbf{b}_r] = [\mathbf{c}', (n+1)\,\mathbf{b}_r],$$

which is of the desired form.

It will be convenient from now on to write a product $\mathbf{xy}$ where either $\mathbf{x}$ or $\mathbf{y}$ may be *empty* in the sense that possibly $\mathbf{xy} = \mathbf{x}$ or $\mathbf{xy} = \mathbf{y}$. This slight abuse of teminology will save much circumlocution. The same terminology may be applied to products of more than two expressions; on the other hand, the idea of a product in which all factors are empty is meaningless *per se*.

Definition 7·5. *With the conditions of lemma* 7·2:

(A) *For any integer* $r \geqslant 0$, *write* $\mathbf{P}_r$ *for the set of expressions of the form*

$$\mathbf{p} = \mathbf{b}_1^{\alpha_1}\mathbf{b}_2^{\alpha_2}\ldots\mathbf{b}_r^{\alpha_r}\mathbf{c}_1\mathbf{c}_2\ldots\mathbf{c}_k,$$

where each α_i *is a non-negative integer,* $k \geqslant 0$ *and each* $\mathbf{c}_i \in \mathbf{X}_r$. *The possibilities* $r = 0$ *and* $k = 0$ *correspond to the possibilities that the products* $\mathbf{b}_1^{\alpha_1}\mathbf{b}_2^{\alpha_2}\ldots\mathbf{b}_r^{\alpha_r}$ *and* $\mathbf{c}_1\mathbf{c}_2\ldots\mathbf{c}_k$ *may be empty. The symbol* $S(\mathbf{p})$ *is defined*

$$S(\mathbf{p}) = \alpha_1\,\mathrm{wt}\,(\mathbf{b}_1) + \alpha_2\,\mathrm{wt}\,(\mathbf{b}_2) + \ldots + \alpha_r\,\mathrm{wt}\,(\mathbf{b}_r) + \mathrm{wt}\,(\mathbf{c}_1) + \mathrm{wt}\,(\mathbf{c}_2) + \ldots + \mathrm{wt}\,(\mathbf{c}_k).$$

For each positive integer w, $\mathbf{P}_r(w)$ *is the set*

$$\mathbf{P}_r(w) = \{\mathbf{p} : \mathbf{p} \in \mathbf{P}_r,\ S(\mathbf{p}) = w\}.$$

(B). *A mapping* $\theta_r : \mathbf{P}_{r-1} \to \mathbf{P}_r$ *is defined for each* $r \geqslant 1$. *Suppose* $p \in \mathbf{P}_{r-1}$. *Then it is of the form*

$$\mathbf{p} = \mathbf{b}_1^{\alpha_1}\mathbf{b}_2^{\alpha_2}\ldots\mathbf{b}_{r-1}^{\alpha_{r-1}}\mathbf{u},$$

where $\mathbf{u}$ *is a (possibly empty) product of commutators from* $\mathbf{X}_{r-1}$. *Now* $\mathbf{u}$ *may or may not contain commutators of the form* $\mathbf{b}_r$ *as factors; in any case it may be written uniquely in the form*

$$\mathbf{u} = \mathbf{b}_r^{\beta_0}\mathbf{a}_1\mathbf{b}_r^{\beta_1}\mathbf{a}_2\mathbf{b}_r^{\beta_2}\ldots\mathbf{a}_m\mathbf{b}_r^{\beta_m},$$

where $m \geqslant 0$, *each* $\beta_i \geqslant 0$ *and, by lemma* 7·3 (i), *each* $\mathbf{a}_i \in \mathbf{X}_{r-1} \cap \mathbf{X}_r$. *Then* $\mathbf{p}\theta_r$ *is defined*

$$\mathbf{p}\theta_r = \mathbf{b}_1^{\alpha_1}\mathbf{b}_2^{\alpha_2}\ldots\mathbf{b}_{r-1}^{\alpha_{r-1}}\mathbf{b}_r^{\beta_0}[\mathbf{a}_1, \beta_1\,\mathbf{b}_r]\,[\mathbf{a}_2, \beta_2\,\mathbf{b}_r]\ldots[\mathbf{a}_m, \beta_m\,\mathbf{b}_r],$$

For $1 \leqslant i \leqslant m$, $[\mathbf{a}_i, \beta_i\,\mathbf{b}_r] \in \mathbf{X}_r$ *by lemma* 7·3 (ii) *and so* $\mathbf{p}\theta_r \in \mathbf{P}_r$ *as promised.*

Lemma 7·4.

(i) θ_r *is a one-to-one mapping of* $\mathbf{P}_{r-1}$ *onto* $\mathbf{P}_r$.

(ii) *If* $\mathbf{p} \in \mathbf{P}_{r-1}$ *then* $S(\mathbf{p}\theta_r) = S(\mathbf{p})$.

Proof. (i) It is sufficient to exhibit a mapping $\theta': \mathbf{P}_r \to \mathbf{P}_{r-1}$ such that $\theta'\theta_r$ is the identity on $\mathbf{P}_r$ and $\theta_r\theta'$ is the identity on $\mathbf{P}_{r-1}$. Suppose then that $\mathbf{p} \in \mathbf{P}_r$. Then it is of the form $\mathbf{p} \in \mathbf{b}_1^{\alpha_1}\mathbf{b}_2^{\alpha_2}\ldots\mathbf{b}_r^{\alpha_r}\mathbf{v}$, where $\mathbf{v}$ is a (possibly empty) product of commutators from $\mathbf{X}_r$. By lemma 7·3 (ii), $\mathbf{v}$ may be written in the form $\mathbf{v} = [\mathbf{a}_1, \beta_1\mathbf{b}_r][\mathbf{a}_2, \beta_2\mathbf{b}_r]\ldots[\mathbf{a}_m, \beta_m\mathbf{b}_r]$, where $m \geqslant 0$, each $\beta_i \geqslant 0$ and each $\mathbf{a}_i \in \mathbf{X}_{r-1} \cap \mathbf{X}_r$. Then defining

$$\mathbf{p}\theta' = \mathbf{b}_1^{\alpha_1}\mathbf{b}_2^{\alpha_2}\ldots\mathbf{b}_{r-1}^{\alpha_{r-1}}.\mathbf{b}_r^{\alpha_r}\mathbf{a}_1\mathbf{b}_r^{\beta_1}\mathbf{a}_2\mathbf{b}_r^{\beta_2}\ldots\mathbf{a}_m\mathbf{b}_r^{\beta_m},$$

it follows that $\mathbf{p}\theta' \in \mathbf{P}_{r-1}$. That $\theta_r\theta'$ and $\theta'\theta_r$ are identities then follows immediately from the definitions of θ_r and θ'.

(ii) This follows by an easy calculation using the definition of $S(\mathbf{p})$ and lemma 7·1 (i).

Corollary. *With the conditions of lemma* 7·2, *for* $r \geqslant 0$ *and* $w \geqslant 1$, $|\mathbf{P}_r(w)| = \tau^w$.

Proof. By the lemma, the restriction of θ_r to $\mathbf{P}_{r-1}(w)$ is a one-to-one mapping of $\mathbf{P}_{r-1}(w)$ on to $\mathbf{P}_r(w)$. Thus $|\mathbf{P}_r(w)| = |\mathbf{P}_0(w)|$. But $\mathbf{P}_0(w)$ is the set of all products $\mathbf{u}$ of elements of $\mathbf{X}_0$ for which $S(\mathbf{u}) = w$. But this is just the set of all expressions of the form $\mathbf{g}_{i_1}\mathbf{g}_{i_2}\ldots\mathbf{g}_{i_w}$, where $i_1, i_2, \ldots, i_w < \tau$. Hence $|\mathbf{P}_0(w)| = \tau^w$.

Lemma 7·5. *Suppose the conditions of lemma* 7·2 *hold. Then the number* m_w *of* $(\leqslant)$*-basic commutators of weight* w *is given recursively by*

(i) $m_1 = \tau$,

(ii) $m_w = \tau^w - \langle m_1, m_2, \ldots, m_{w-1}\rangle \quad (w > 1)$,

where the integer $\langle m_1, m_2, \ldots, m_k\rangle$ *is the number of functions* f *from the set*

$$K_k = \{(i,j): 1 \leqslant i \leqslant k,\ 1 \leqslant j \leqslant m_i\}$$

into the non-negative integers satisfying

$$\sum_{(i,j)\in K_k} i.f(i,j) = k+1.$$

Proof. By induction over w. The result $m_1 = \tau$ is already known. Now suppose the result is true for all weights up to $w-1$.

By lemma 7·2, there exists an integer $N(=n_{w+1})$ such that $i \geqslant N \Rightarrow \mathrm{wt}(\mathbf{b}_i) > w$. Consider $\mathbf{P}_N(w)$ An element $\mathbf{p} \in \mathbf{P}_N(w)$ is of the form $\mathbf{p} = \mathbf{b}_1^{\alpha_1}\mathbf{b}_2^{\alpha_2}\ldots\mathbf{b}_N^{\alpha_N}\mathbf{u}$ where $\mathbf{u}$ is a (possibly empty) product of elements of $\mathbf{X}_N$. But $\mathbf{u}$ is in fact empty, for otherwise $\mathbf{u} = \mathbf{u}'\mathbf{b}_h$ where $\mathbf{b}_h \in \mathbf{X}_N$ and then, since $[\mathbf{b}_h, \mathbf{b}_N]$ is $(\leqslant)$-basic, $h \geqslant N$ so that $S(\mathbf{p}) \geqslant S(\mathbf{b}_h) > w$. Thus each element of $\mathbf{P}_N(w)$ is of the form $\mathbf{p} = \mathbf{b}_1^{\alpha_1}\mathbf{b}_2^{\alpha_2}\ldots\mathbf{b}_N^{\alpha_N}$.

Now all the $(\leqslant)$-basic commutators of weight $\leqslant w$ appear in $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N\}$ together possibly with some of higher weight (for $\leqslant$ does not necessarily preserve weight). However, since $S(\mathbf{p}) = w$, any commutator $\mathbf{b}_i$ $(1 \leqslant i \leqslant N)$ of weight $> w$ must have exponent $\alpha_i = 0$ in the given expression for $\mathbf{p}$. Thus $\mathbf{p}$ is defined uniquely by the powers α_i of commutators of weight $\leqslant w$ in that expression.

Suppose that the $(\leqslant)$-basic commutators of weight $\leqslant w$ are re-indexed as follows:

$$\begin{matrix} \mathbf{b}_{(1,1)} & \mathbf{b}_{(1,2)} & \cdots & \mathbf{b}_{(1,m_1)} \\ \mathbf{b}_{(2,1)} & \mathbf{b}_{(2,2)} & \cdots & \mathbf{b}_{(2,m_2)} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{b}_{(w,1)} & \mathbf{b}_{(w,2)} & \cdots & \mathbf{b}_{(w,m_w)} \end{matrix}$$

where, for each c $(1 \leqslant c \leqslant w)$, $\{\mathbf{b}_{(c,1)}, \mathbf{b}_{(c,2)}, \ldots, \mathbf{b}_{(c,m_c)}\}$ is the set of $(\leqslant)$-basic commutators of weight c (in any order). Thus the set of $(\leqslant)$-basic commutators of weight $\leqslant w$ has been indexed by the set K_w defined in the statement of the lemma. The order in which the commutators are written down in this array is not, of course, their order under the B-order $\leqslant$.

Let $\mathbf{Q}$ be the set of all $\mathbf{p} \in \mathbf{P}_N(w)$ in which some commutator of weight w has non-zero exponent. Then, since $S(\mathbf{p}) = w$, $\mathbf{Q}$ is exactly the set of $(\leqslant)$-basic commutators of weight w and so $|\mathbf{Q}| = m_w$.

But now, if to each $\mathbf{p} \in \mathbf{P}_N(w) - \mathbf{Q}$ a function $f_\mathbf{p}$ from the set K_{w-1} into the non-negative integers is defined by setting $f_\mathbf{p}(i,j)$ to be the power of $\mathbf{b}_{(i,j)}$ appearing in the expression for $\mathbf{p}$, it follows that $f_\mathbf{p}$ is uniquely determined by $\mathbf{p}$ and that

$$\sum_{(i,j) \in K_{w-1}} i \,.\, f_\mathbf{p}(i,j) = S(\mathbf{p}) = w$$

and conversely that any such function uniquely defines an element of $\mathbf{P}_N(w) - \mathbf{Q}$. Thus

$$|\mathbf{P}_N(w) - Q| = \langle m_1, m_2, \ldots, m_{w-1} \rangle$$

and this, together with the fact just proved that $|\mathbf{Q}| = m_w$ and the corollary to lemma 7·4, proves this lemma.

Theorem 7·1. *Let $\sigma: \mathbf{A} \to W$ be any shape and suppose that the number τ of generators of $\mathbf{A}$ is finite. Then the number m_w of W-basic commutators of weight w is given by Witt's formula:*

$$m_w = \frac{1}{w} \sum_{r|w} \mu(r)\, \tau^{w/r},$$

where μ is the Möbius function.

Proof. Let $\leqslant$ be the W-ordering of $\mathbf{C}$ and w be given. Now construct a new full-order $\leqslant'$ on $\mathbf{C}$ by specifying that $\leqslant'$ coincides with $\leqslant$ on the set of commutators of weight $\leqslant w$ (that is, if $\mathbf{a}$ and $\mathbf{b}$ are commutators of weight $\leqslant w$, then $\mathbf{a} \leqslant' \mathbf{b} \Leftrightarrow \mathbf{a} \leqslant \mathbf{b}$) and then extending $\leqslant'$ in any way that preserves weight to the remainder of $\mathbf{C}$ (so that, if $\mathbf{a}$ and $\mathbf{b}$ are commutators such that wt $(\mathbf{a}) <$ wt $(\mathbf{b}) > w$ then $\mathbf{a} <' \mathbf{b}$).

Now, since τ is finite, the number of commutators of each weight is finite, so $\leqslant'$ is of order-type ω. Further, $\leqslant'$ is clearly a B-order (see definition 7·1). Thus lemma 7·5 obtains and the number m'_w of $(\leqslant')$-basic commutators of each weight c is given by the recursive definition

$$m'_1 = \tau, \quad m'_c = \tau^c - \langle m'_1, m'_2, \ldots, m'_{c-1} \rangle.$$

But $\leqslant'$ was chosen to coincide with $\leqslant$ on the set of all commutators of weight $\leqslant w$ so, for each $c \leqslant w$, the collection of $(\leqslant')$-basic commutators of weight c is the same as the collection of $(\leqslant)$-basic commutators of weight c and this in turn is just the collection of W-basic commutators of weight c (compare definitions 7·1 (B) and 6·2 (A)). Thus, for each $c \leqslant w$, $m'_c = m_c$. This proves the recursive formula

$$m_1 = \tau, \quad m_c = \tau^c - \langle m_1, m_2, \ldots, m_{c-1} \rangle \quad (c \leqslant w).$$

But this formula depends only upon τ and w and not on the particular shape range W chosen: in particular m_w is the same as the number of N^--basic commutators of weight w. But the N^--ordering of $\mathbf{C}$ preserves weight, so these are just the commutators of weight w which are basic in the conventional sense and so m_w is given by Witt's formula.

Chapter II. The collecting processes

The collecting processes to be described in this chapter differ in two important respects from the process used in the conventional theory. First, the processes described here are defined in terms of a particular shape range W, the object being to convert an arbitrary expression in **A** to a W-basic one; in this sense these processes are more general than the conventional one. Secondly, while the conventional processes involve an initial expansion of the expression to be collected into a product of generators and their inverses, an expression of shape 1, followed by collection into products of basic commutators of successively higher weights, the processes described here involve no such initial expansion: they proceed through a sequence of expressions of non-decreasing shape and so certain properties of commutators which may be expressed in terms of their shape are preserved.

It is well known that calculations performed in the 'bottom' of a nilpotent group, that is in $\gamma_c(G)$ when G is nilpotent of class c, usually have a particularly simple form. Accordingly it will be advantageous to describe first a collecting process which operates in the 'bottom' of a group, however here the 'bottom' may also mean $W_\alpha(G)$ when $G \in \mathfrak{W}_{\alpha+1}$. This is the *special* process. Following this a *general* process will be described which can operate either anywhere in $W_\alpha(G)$ when $G \in \mathfrak{W}_\beta$ under certain restrictions on α and β or else anywhere in a nilpotent group.

For the remainder of this chapter it will be assumed that a fixed algebra **A** of expressions and a fixed shape range W are under consideration and all definitions will be made in terms of these.

8. *The special process*

Definition 8·1. (A) *Let* $\mathbf{x}, \mathbf{y} \in \mathbf{A}$. *Then write* $d: \mathbf{x} \to \mathbf{y}$ *if* $\mathbf{x}$ *and* $\mathbf{y}$ *are any of the following forms*:

(i) $\mathbf{x} = \mathbf{ab}, \quad \mathbf{y} = \mathbf{ba}$.

(ii) $\mathbf{x} = (\mathbf{ab})^{-1}, \quad \mathbf{y} = \mathbf{b}^{-1}\mathbf{a}^{-1}$.

(iii) $\mathbf{x} = (\mathbf{a}^{-1})^{-1}, \quad \mathbf{y} = \mathbf{a}$.

(iv) $\mathbf{x} = \mathbf{a}^{-1}\mathbf{a}$ or $\mathbf{a}\mathbf{a}^{-1}, \quad \mathbf{y} = \mathbf{1}$.

(v) $\mathbf{x} = \mathbf{a1}$ or $\mathbf{1a}, \quad \mathbf{y} = \mathbf{a}$.

(vi) $\mathbf{x} = \mathbf{1}^{-1}, \quad \mathbf{y} = \mathbf{1}$.

(vii) $\mathbf{x} = [\mathbf{a}, \mathbf{a}], \quad \mathbf{y} = \mathbf{1}$.

(viii) $\mathbf{x} = [\mathbf{a}, \mathbf{1}]$ or $[\mathbf{1}, \mathbf{a}], \quad \mathbf{y} = \mathbf{1}$.

(ix) $\mathbf{x} = [\mathbf{a}^{-1}, \mathbf{b}]$ or $[\mathbf{a}, \mathbf{b}^{-1}], \quad \mathbf{y} = [\mathbf{a}, \mathbf{b}]^{-1}$.

(x) $\mathbf{x} = [\mathbf{ab}, \mathbf{c}], \quad \mathbf{y} = [\mathbf{a}, \mathbf{c}]\,[\mathbf{b}, \mathbf{c}]$ *or* $\mathbf{x} = [\mathbf{a}, \mathbf{bc}], \quad \mathbf{y} = [\mathbf{a}, \mathbf{c}]\,[\mathbf{a}, \mathbf{b}]$.

(xi) $\mathbf{x} = [\mathbf{a}, \mathbf{b}], \quad \mathbf{y} = [\mathbf{b}, \mathbf{a}]^{-1}$ *provided* $\mathbf{a}, \mathbf{b}$ *are commutators and* $\mathbf{a} < \mathbf{b}$.

(xii) $\mathbf{x} = [\mathbf{c}, \mathbf{b}, \mathbf{a}], \quad \mathbf{y} = [\mathbf{b}, \mathbf{a}, \mathbf{c}]^{-1}\,[\mathbf{c}, \mathbf{a}, \mathbf{b}]$ *provided* $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{c}$ *are commutators and* $\mathbf{a} < \mathbf{b} < \mathbf{c}$.

The notation is extended to larger expressions by recursion over their height:

(xiii) *If* $d: \mathbf{a}_1 \to \mathbf{a}_2$ *then* $d: \mathbf{a}_1^{-1} \to \mathbf{a}_2^{-1}$ *and, for any* $\mathbf{b} \in \mathbf{A}$, $d: \mathbf{a}_1\mathbf{b} \to \mathbf{a}_2\mathbf{b}$, $d: \mathbf{b}\mathbf{a}_1 \to \mathbf{b}\mathbf{a}_2$, $d: [\mathbf{a}_1, \mathbf{b}] \to [\mathbf{a}_2, \mathbf{b}]$ *and* $d: [\mathbf{b}, \mathbf{a}_1] \to [\mathbf{b}, \mathbf{a}_2]$.

(B) *Write* $D: \mathbf{x} \to \mathbf{y}$ *if there exists a finite sequence* $(\mathbf{u}_i)_{i=0}^{k}$ $(k \geqslant 0)$ *of expressions such that* $\mathbf{u}_0 = \mathbf{x}$, $\mathbf{u}_k = \mathbf{y}$ *and* $d: \mathbf{u}_{i-1} \to \mathbf{u}_i$ $(1 \leqslant i \leqslant k)$.

The relation $D: \mathbf{x} \to \mathbf{y}$ is clearly reflexive and transitive, that is, it is a pre-order. It also follows from the definition that Part (xiii) holds just as well for D as for d.

Definition 8·2. *A* product of commutators *of* length l *is defined*:

(i) $\mathbf{1}$ *is a product of commutators of length* 1, *as are* $\mathbf{c}$ *and* $\mathbf{c}^{-1}$ *where* $\mathbf{c}$ *is a commutator.*

(ii) *If* $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *are products of commutators of lengths* l_1 *and* l_2 *respectively then* $\mathbf{x}_1\mathbf{x}_2$ *is a product of commutators of length* l_1+l_2.

Clearly a product of commutators $\mathbf{x}$ has a property $\mathscr{P}$ essentially (definition 1·4) if and only if each of the commutators which are factors of $\mathbf{x}$ has $\mathscr{P}$.

Definition 8·3. *A relation* $\leqslant^0$ *is defined on the set* $\mathbf{C}$ *of commutators by*: $\mathbf{a} \leqslant^0 \mathbf{b}$ *if and only if*

(i) $\sigma(\mathbf{a}) > \sigma(\mathbf{b})$ *or*

(ii) $\sigma(\mathbf{a}) = \sigma(\mathbf{b})$ *and* $\mathbf{a} \leqslant \mathbf{b}$.

Clearly $\leqslant^0$ is a full order but not a well-order. The relation $<^0$ is defined in the obvious way.

Lemma 8·1. *If* $\mathbf{a}_1$, $\mathbf{a}_2$ *and* $\mathbf{b}$ *are commutators and* $\mathbf{a}_1 < \mathbf{a}_2$ *then*

$$[\mathbf{a}_1, \mathbf{b}] \text{ and } [\mathbf{b}, \mathbf{a}_1] < [\mathbf{a}_2, \mathbf{b}] \text{ and } [\mathbf{b}, \mathbf{a}_2].$$

Proof is only given that $[\mathbf{a}_1, \mathbf{b}] < [\mathbf{a}_2, \mathbf{b}]$, the proofs of the other three inequalities being similar. Since $\mathbf{a}_1 < \mathbf{a}_2$, $\sigma(\mathbf{a}_1) \leqslant \sigma(\mathbf{a}_2)$. If $\sigma(\mathbf{a}_1) < \sigma(\mathbf{a}_2)$ then $\sigma([\mathbf{a}_1, \mathbf{b}]) < \sigma([\mathbf{a}_2, \mathbf{b}])$ and the lemma is true. Otherwise $\sigma(\mathbf{a}_1) = \sigma(\mathbf{a}_2)$ so that $\sigma([\mathbf{a}_1, \mathbf{b}]) = \sigma([\mathbf{a}_2, \mathbf{b}])$ and there are two possibilities: first, if $\mathbf{b} < \mathbf{a}_2$ then

$$\mathrm{ld}\,([\mathbf{a}_1, \mathbf{b}]) = \mathbf{a}_1 \quad\text{or}\quad \mathbf{b} < \mathbf{a}_2 = \mathrm{ld}\,([\mathbf{a}_2, \mathbf{b}]),$$

and then $[\mathbf{a}_1, \mathbf{b}] < [\mathbf{a}_2, \mathbf{b}]$ and secondly if $\mathbf{b} \geqslant \mathbf{a}_2$, then

$$\mathrm{ld}\,([\mathbf{a}_1, \mathbf{b}]) = \mathbf{b} = \mathrm{ld}\,([\mathbf{a}_2, \mathbf{b}]) \quad\text{and}\quad \mathrm{tr}\,([\mathbf{a}_1, \mathbf{b}]) = \mathbf{a}_1 < \mathbf{a}_2 = \mathrm{tr}\,([\mathbf{a}_2, \mathbf{b}])$$

so again $[\mathbf{a}_1, \mathbf{b}] < [\mathbf{a}_2, \mathbf{b}]$.

Corollary. *If* $\leqslant$, $<^0$ *or* $\leqslant^0$ *are substituted for* $<$ *in the lemma, it is still true.*

Lemma 8·2. *Let* $\mathbf{x}$ *be an expression. Then there exists a product* $\mathbf{y}$ *of commutators such that* $D\colon \mathbf{x} \rightarrow \mathbf{y}$.

Proof. If $\mathbf{z}$ is a product of commutators then there exists a product of commutators $\mathbf{y}$ such that $D\colon \mathbf{z}^{-1} \rightarrow \mathbf{y}$: this follows by an easy induction over the length of $\mathbf{z}$ using only parts (A) (ii), (iii) and (vi) and part (B) of definition 8·1. If $\mathbf{z}_1$ and $\mathbf{z}_2$ are two products of commutators then there exists a product of commutators $\mathbf{y}$ such that $D\colon [\mathbf{z}_1, \mathbf{z}_2] \rightarrow \mathbf{y}$: this also follows by induction over the lengths of $\mathbf{z}_1$ and $\mathbf{z}_2$ using only parts (A) (viii) and (ix) and part (B) of definition 8·1 and the first remark in this proof. The lemma now follows by an easy induction over the height of $\mathbf{x}$.

Lemma 8·3. *Suppose* $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{c}$ *are commutators and* $\mathbf{a} < \mathbf{b} < \mathbf{c}$. *Then*

$$[\mathbf{c}, \mathbf{a}, \mathbf{b}] <^0 [\mathbf{c}, \mathbf{b}, \mathbf{a}] \quad\text{and}\quad [\mathbf{b}, \mathbf{a}, \mathbf{c}] <^0 [\mathbf{c}, \mathbf{b}, \mathbf{a}].$$

Proof. Since $\mathbf{a} < \mathbf{b} < \mathbf{c}$, $\sigma(\mathbf{a}) \leqslant \sigma(\mathbf{b}) \leqslant \sigma(\mathbf{c})$ so $\sigma([\mathbf{c}, \mathbf{a}, \mathbf{b}]) = \sigma([\mathbf{c}, \mathbf{b}, \mathbf{a}])$ by definition 2·1 (v). Further $[\mathbf{c}, \mathbf{a}] > \mathbf{c} > \mathbf{b}$ and $[\mathbf{c}, \mathbf{b}] > \mathbf{c} > \mathbf{a}$ so

$$\mathrm{ld}\,([\mathbf{c}, \mathbf{a}, \mathbf{b}]) = [\mathbf{c}, \mathbf{a}] \quad\text{and}\quad \mathrm{ld}\,([\mathbf{c}, \mathbf{b}, \mathbf{a}]) = [\mathbf{c}, \mathbf{b}].$$

But $[\mathbf{c}, \mathbf{a}] < [\mathbf{c}, \mathbf{b}]$ since $\mathbf{a} < \mathbf{b}$ (lemma 8·1) so $[\mathbf{c}, \mathbf{a}, \mathbf{b}] < [\mathbf{c}, \mathbf{b}, \mathbf{a}]$ (definition 6·1 (iii*a*)). Then, since their shapes are the same, $[\mathbf{c}, \mathbf{a}, \mathbf{b}] <^0 [\mathbf{c}, \mathbf{b}, \mathbf{a}]$.

By definition 2·1 (v) again, $\sigma([\mathbf{b},\mathbf{a},\mathbf{c}]) \geqslant \sigma([\mathbf{c},\mathbf{b},\mathbf{a}])$. If $\sigma([\mathbf{b},\mathbf{a},\mathbf{c}]) > \sigma([\mathbf{c},\mathbf{b},\mathbf{a}])$ then $[\mathbf{b},\mathbf{a},\mathbf{c}] <^0 [\mathbf{c},\mathbf{b},\mathbf{a}]$ immediately. If $\sigma([\mathbf{b},\mathbf{a},\mathbf{c}]) = \sigma([\mathbf{c},\mathbf{b},\mathbf{a}])$ it is sufficient to prove that $\mathrm{ld}([\mathbf{b},\mathbf{a},\mathbf{c}]) < \mathrm{ld}([\mathbf{c},\mathbf{b},\mathbf{a}])$. But $\mathrm{ld}([\mathbf{b},\mathbf{a},\mathbf{c}])$ is either $[\mathbf{b},\mathbf{a}]$ or $\mathbf{c}$, both of which are $< [\mathbf{c},\mathbf{b}] = \mathrm{ld}([\mathbf{c},\mathbf{b},\mathbf{a}])$ so the lemma is true.

COROLLARY. *Suppose* $\mathbf{a}$ *is a non-basic commutator* (*a commutator which is not a basic one*). *Then* $D\colon \mathbf{a} \to \mathbf{b}$, *where* $\mathbf{b}$ *is one of the forms*:

(i) $\mathbf{b} = \mathbf{1}$, (ii) $\mathbf{b} = \mathbf{c}^{\epsilon}$ *where* $\mathbf{c} \in \mathbf{C}$, $\epsilon = \pm 1$, $\mathbf{c} <^0 \mathbf{a}$ *and* $\mathrm{wt}(\mathbf{c}) = \mathrm{wt}(\mathbf{a})$,

(iii) $\mathbf{b} = \mathbf{c}_1^{\epsilon_1}\mathbf{c}_2^{\epsilon_2}$ *where* $\mathbf{c}_1$ *and* $\mathbf{c}_2 \in \mathbf{C}$, ϵ_1 *and* $\epsilon_2 = \pm 1$, $\mathbf{c}_1$ *and* $\mathbf{c}_2 <^0 \mathbf{a}$ *and* $\mathrm{wt}(\mathbf{c}_1) = \mathrm{wt}(\mathbf{c}_2) = \mathrm{wt}(\mathbf{a})$.

Proof. The argument is by induction over the weight of $\mathbf{a}$. If $\mathrm{wt}(\mathbf{a}) = 1$ then $\mathbf{a}$ is basic and the lemma is vacuously true. Now suppose that $\mathrm{wt}(\mathbf{a}) > 1$ and the lemma is true for smaller weights. Since $\mathbf{a}$ is non-basic, at least one of the conditions of definition 6·2 (A) must fail. These are treated separately. Write $\mathbf{a} = [\mathbf{a}_1, \mathbf{a}_2]$.

$\mathbf{a}_1$ *is non-basic*. Then by the inductive hypothesis, $D\colon \mathbf{a}_1 \to \mathbf{b}_1$ and so $D\colon \mathbf{a} \to [\mathbf{b}_1, \mathbf{a}_2]$ where $\mathbf{b}_1$ has one of the forms given above. If $\mathbf{b}_1 = \mathbf{1}$ then $D\colon \mathbf{a} \to \mathbf{1}$ by definition 8·1 (A) (viii). If $\mathbf{b}_1 = \mathbf{c}^{\epsilon}$ where $\mathbf{c} \in \mathbf{C}$, $\epsilon = \pm 1$, $\mathbf{c} <^0 \mathbf{a}_1$ and $\mathrm{wt}(\mathbf{c}) = \mathrm{wt}(\mathbf{a}_1)$ then $D\colon \mathbf{a} \to [\mathbf{c}, \mathbf{a}_2]^{\epsilon}$ by definition 8·1 (A) (ix), $[\mathbf{c}, \mathbf{a}_2] \in \mathbf{C}$, $[\mathbf{c}, \mathbf{a}_2] <^0 \mathbf{a}$ by the corollary to lemma 8·1 and $\mathrm{wt}([\mathbf{c}, \mathbf{a}_2]) = \mathrm{wt}(\mathbf{a})$. The argument when $\mathbf{b}_1 = \mathbf{c}_1^{\epsilon_1}\mathbf{c}_2^{\epsilon_2}$ is similar.

$\mathbf{a}_2$ *is non-basic*. The argument is the same as that just given.

$\mathbf{a}_1 \leqslant \mathbf{a}_2$. Then either $\mathbf{a}_1 = \mathbf{a}_2$ in which case $D\colon \mathbf{a} \to \mathbf{1}$ by definition 8·1 (A) (vii) or else $\mathbf{a}_1 < \mathbf{a}_2$, in which case $D\colon \mathbf{a} \to [\mathbf{a}_2, \mathbf{a}_1]^{-1}$ by definition 8·1 (A) (xi). But then

$$\sigma([\mathbf{a}_2, \mathbf{a}_1]) = \sigma(\mathbf{a}), \quad \mathrm{ld}([\mathbf{a}_2, \mathbf{a}_1]) = \mathbf{a}_2 = \mathrm{ld}(\mathbf{a}),$$
$$\mathrm{tr}([\mathbf{a}_2, \mathbf{a}_1]) = \mathbf{a}_1 = \mathrm{tr}(\mathbf{a}) \quad \text{and} \quad \mathbf{a}_1 < \mathbf{a}_2.$$

Thus $[\mathbf{a}_2, \mathbf{a}_1] < \mathbf{a}$ by definition 6·1 (iii*c*) and so $[\mathbf{a}_2, \mathbf{a}_1] <^0 \mathbf{a}$. Clearly $\mathrm{wt}([\mathbf{a}_2, \mathbf{a}_1]) = \mathrm{wt}(\mathbf{a})$.

$\mathbf{a}_1 = [\mathbf{a}_{11}, \mathbf{a}_{12}]$ *and* $\mathbf{a}_{12} > \mathbf{a}_2$. By virtue of the first case considered, it may be assumed that $\mathbf{a}_1$ is basic and thus that $\mathbf{a}_{11} > \mathbf{a}_{12}$. Then $\mathbf{a}_{11} > \mathbf{a}_{12} > \mathbf{a}_2$ and $\mathbf{a} = [\mathbf{a}_{11}, \mathbf{a}_{12}, \mathbf{a}_2]$ so

$$D\colon \mathbf{a} \to [\mathbf{a}_{12}, \mathbf{a}_2, \mathbf{a}_{11}]^{-1}[\mathbf{a}_{11}, \mathbf{a}_2, \mathbf{a}_{12}]$$

by definition 8·1 (A) (xii). But now $[\mathbf{a}_{12}, \mathbf{a}_2, \mathbf{a}_{11}] <^0 \mathbf{a}$ and $[\mathbf{a}_{11}, \mathbf{a}_2, \mathbf{a}_{12}] <^0 \mathbf{a}$ by the lemma. Again it is clear that $\mathrm{wt}([\mathbf{a}_{12}, \mathbf{a}_2, \mathbf{a}_{11}]) = \mathrm{wt}([\mathbf{a}_{11}, \mathbf{a}_2, \mathbf{a}_{12}) = \mathrm{wt}(\mathbf{a})$.

LEMMA 8·4. *For any positive integer* w, *the set of commutators of weight not exceeding* w *is* (*fully*) *well-ordered by* $\leqslant^0$.

Proof. Consider the set

$$\{\sigma(\mathbf{x}) : \mathbf{x} \in \mathbf{C},\ \mathrm{wt}(\mathbf{x}) \leqslant w\}$$

of all possible shapes that a commutator of weight $\leqslant w$ may have. It follows by an easy induction over w that this set is finite. But for any $\alpha \in W$, the order $\leqslant^0$ coincides with the W-ordering $\leqslant$ on the set of all commutators of shape α and so well-orders that set. It follows that, as far as $\leqslant^0$ is concerned, the set of all commutators of weight $\leqslant w$ is the union of a finite number of well-ordered sets. The lemma follows.

LEMMA 8·5. *If* $D\colon \mathbf{x} \to \mathbf{y}$ *then* $\Sigma(\mathbf{x}) \preccurlyeq \Sigma(\mathbf{y})$, *these sets being ordered according to definition* 2·2 (i), $\hat{\sigma}(\mathbf{x}) \preccurlyeq \hat{\sigma}(\mathbf{y})$, $\sigma(\mathbf{x}) \leqslant \sigma(\mathbf{y})$ *and* $\mathrm{wt}(\mathbf{x}) \leqslant \mathrm{wt}(\mathbf{y})$.

Proof. It will be shown that $\Sigma(\mathbf{x}) \preccurlyeq \Sigma(\mathbf{y})$; the proofs of the corresponding results for the coarse and fine shapes and weight are similar.

Check the various parts of definition 8·1: for parts (A) (i) to (vi) and (viii) to (xi), $\Sigma(\mathbf{x}) = \Sigma(\mathbf{y})$, for part (vii), $\Sigma(\mathbf{y}) = \varnothing \geqslant \Sigma(\mathbf{x})$ and for part (xii)

$$\Sigma(\mathbf{x}) = \{\hat{\sigma}(\mathbf{c})+\hat{\sigma}(\mathbf{b})+\hat{\sigma}(\mathbf{a})\}$$

and
$$\Sigma(\mathbf{y}) = \{\hat{\sigma}(\mathbf{b})+\hat{\sigma}(\mathbf{a})+\hat{\sigma}(\mathbf{c}),\ \hat{\sigma}(\mathbf{c})+\hat{\sigma}(\mathbf{a})+\hat{\sigma}(\mathbf{c})\}.$$

But then
$$\hat{\sigma}(\mathbf{c})+\hat{\sigma}(\mathbf{b})+\hat{\sigma}(\mathbf{a}) = \hat{\sigma}(\mathbf{c})+\hat{\sigma}(\mathbf{a})+\hat{\sigma}(\mathbf{c}) \leqslant \hat{\sigma}(\mathbf{b})+\hat{\sigma}(\mathbf{a})+\hat{\sigma}(\mathbf{c})$$

by definition 2·1 (v), using the fact that **a**, **b** and **c** are all commutators so that

$$\hat{\sigma}(\mathbf{a}) = \sigma(\mathbf{a}),\quad \hat{\sigma}(\mathbf{b}) = \sigma(\mathbf{b})\quad \text{and}\quad \hat{\sigma}(\mathbf{c}) = \sigma(\mathbf{c}).$$

For part (A) (xiii) the result follows immediately from the definition of Σ and for part (B) the result follows immediately from the definition of Σ and for part (B) the result follows from transitivity of $\leqslant$.

The following theorem and its corollaries are out of logical order in this study. They are placed here because they summarize the properties of the special collecting process and provide a motivation for the definitions of this section. The proof requires the results of lemmas 9·1 and 9·2, so that the theorem strictly should be stated and proved immediately following the latter lemma.

Theorem 8·1. *For any expression* $\mathbf{x} \in \mathbf{A}$ *there exists* $\mathbf{y} \in \mathbf{A}$ *such that*

(i) $\mathbf{y}$ *is a basic expression,*
(ii) $D\colon \mathbf{x} \to \mathbf{y}$,
(iii) $\sigma(\mathbf{x}) \leqslant \sigma(\mathbf{y})$, $\hat{\sigma}(\mathbf{x}) \leqslant \hat{\sigma}(\mathbf{y})$, $\Sigma(\mathbf{x}) \leqslant \Sigma(\mathbf{y})$ *and* $\mathrm{wt}(\mathbf{x}) \leqslant \mathrm{wt}(\mathbf{y})$, *and*
(iv) *there exists an expression* $\mathbf{yu}$ ($\mathbf{u}$ *possibly empty*) *such that*

$$\sigma(\mathbf{u}) \geqslant \sigma(\mathbf{x})+1,\quad \hat{\sigma}(\mathbf{u}) \geqslant \hat{\sigma}(\mathbf{x})+1,\quad \Sigma(\mathbf{u}) \geqslant \Sigma(\mathbf{x})+\{1\}\quad \textit{and}\quad \mathrm{wt}(\mathbf{u}) \geqslant \mathrm{wt}(\mathbf{x})+1$$

when $\mathbf{u}$ *exists and, for any description* $\rho\colon \mathbf{A} \to G$ *of a group* G, $\mathbf{x}\rho = (\mathbf{yu})\rho$.

Proof. By lemma 8·2 there exists a product of commutators $\mathbf{z}_1$ such that $D\colon \mathbf{x} \to \mathbf{z}_1$. Since the length of $\mathbf{z}_1$ is necessarily finite it follows that there exists an integer w such that $\mathbf{z}_1$ is essentially of weight $\leqslant w$, that is, $\mathbf{z}_1$ is a product of commutators each of weight $\leqslant w$.

It is now shown that $D\colon \mathbf{z}_1 \to \mathbf{z}_2$ where $\mathbf{z}_2$ is essentially basic (a product of basic commutators). If $\mathbf{z}_1$ is itself a product of basic commutators then this is true with $\mathbf{z}_2 = \mathbf{z}_1$. Otherwise, again since the length of $\mathbf{z}_1$ is finite, there exists a non-basic commutator **a** which is maximum under the order $\leqslant^0$ among those factors of $\mathbf{z}_1$ (members of $\Xi(\mathbf{z}_1)$) which are non-basic. This commutator may appear more than once but in any case $\mathbf{z}_1$ may be written in the form

$$\mathbf{z}_1 = \mathbf{v}_0\mathbf{a}^{\epsilon_1}\mathbf{v}_1\mathbf{a}^{\epsilon_2}\mathbf{v}_2\ldots\mathbf{a}^{\epsilon_k}\mathbf{v}_k,$$

where $k \geqslant 1$, each $\epsilon_i = \pm 1$ and each $\mathbf{v}_i$ is a (possibly empty) product of commutators which are either basic or $<^0 \mathbf{a}$. But, by the corollary to lemma 8·3, $D\colon \mathbf{a} \to \mathbf{b}$ where **b** is a product of commutators each of which is $<^0$ **a** and of the same weight as **a**. Thus

$$D\colon \mathbf{z}_1 \to \mathbf{v}_0\mathbf{b}^{\epsilon_1}\mathbf{v}_1\mathbf{b}^{\epsilon_2}\mathbf{v}_2\ldots\mathbf{b}^{\epsilon_k}\mathbf{v}_k,$$

and then application of definition 8·1 (ii) converts this into a product of commutators each of which is of weight $\leqslant w$ and is either basic or $<^0$ **a**. Thus the maximum (under $\leqslant^0$) non-basic factor of $\mathbf{z}_1$ has been replaced by an earlier one. But then by lemma 8·4 this can only be done a finite number of times and eventually $D: \mathbf{z}_1 \to \mathbf{z}_2$ where $\mathbf{z}_2$ is a product of basic commutators.

But then $D: \mathbf{z}_2 \to \mathbf{y}$ where $\mathbf{y}$ is a basic expression: this follows immediately from definition 8·1 (A) (i), (iv) and (v).

Parts (i) and (ii) of the theorem are thus proved. Part (iii) now follows from lemma 8·5.

For part (iv), lemmas 9·1 and 9·2 must be invoked. By lemma 9·1, $E: \mathbf{x} \to \mathbf{yu}$ where $\mathbf{u}$ has the required properties and then, by lemma 9·2, $\mathbf{x}\rho = (\mathbf{yu})\rho$.

Corollary 1. *Suppose Φ is a subset of W, G is a group of the variety $\hat{\mathfrak{W}}_{\Phi+\{1\}}$ and $x \in \hat{W}_\Phi(G)$. If $\rho: \mathbf{A} \to G$ is any description of G then there exists $\mathbf{x} \in \mathbf{A}$ such that $\mathbf{x} \in \hat{\mathbf{B}}_{\Phi+\{1\}} \cap \hat{\mathbf{W}}_\Phi$ (see definitions 6·2 (B) and 3·1) and $\mathbf{x}\rho = x$.*

Proof. Since $x \in \hat{W}_\Phi(G)$, there exists $\mathbf{x}' \in \mathbf{A}$ such that $\mathbf{x}'\rho = x$ and $\Sigma(\mathbf{x}') \geqslant \Phi$. By the theorem, $D: \mathbf{x}' \to \mathbf{x}''$ where $\mathbf{x}''$ is a basic expression and $\Sigma(\mathbf{x}'') \geqslant \Phi$ and $x = \mathbf{x}\rho = (\mathbf{x}''\mathbf{u})\rho$ where $\mathbf{u}$ is possibly empty, but if it exists, $\Sigma(\mathbf{u}) \geqslant \Sigma(\mathbf{x}') + \{1\} \geqslant \Phi + \{1\}$. Then, since $G \in \hat{\mathfrak{W}}_{\Phi+\{1\}}$, $\mathbf{u}\rho = 1$ and $\mathbf{x}''\rho = x$. If $\mathbf{x}'' = \mathbf{1}$ or $\Sigma(\mathbf{x}'') \geqslant \Phi + \{1\}$ then $x = 1$ and the result is true with $\mathbf{x} = \mathbf{1}$. Otherwise $\mathbf{x}''$ is a basic expression of the form

$$\mathbf{x}'' = \mathbf{b}_1^{\alpha_1}\mathbf{b}_2^{\alpha_2}\ldots\mathbf{b}_k^{\alpha_k} \quad (k \geqslant 1),$$

where at least one of the commutators $\mathbf{b}_i^{\alpha_i}$ has the property $\Sigma(\mathbf{b}_i) \not\geqslant \Phi + \{1\}$. Then the result is true with

$$\mathbf{x} = \mathbf{b}_{i_1}^{\alpha_{i_1}}\mathbf{b}_{i_2}^{\alpha_{i_2}}\ldots\mathbf{b}_{i_m}^{\alpha_{i_m}}$$

where $i_1, i_2, \ldots, i_m$ is the subsequence of $1, 2, \ldots, k$ for which the corresponding commutator is of shape set $\not\geqslant \Phi + \{1\}$.

Corollary 2. *Suppose $\phi \in W$, G is a group of the variety $\hat{\mathfrak{W}}_{\phi+1}$ and $x \in \hat{W}_\phi(G)$. If $\rho: \mathbf{A} \to G$ is any description of G then there exists $\mathbf{x} \in \mathbf{A}$ such that $\mathbf{x} \in \hat{\mathbf{B}}_{\phi+1} \cap \hat{\mathbf{W}}_\phi$ and $\mathbf{x}\rho = x$.*

Proof. This follows from corollary 1 by writing $\Phi = \{\phi\}$.

Corollary 3. *Suppose $\phi \in W$, G is a group of the variety $\mathfrak{W}_{\phi+1}$ and $x \in W_\phi(G)$. If $\rho: \mathbf{A} \to G$ is any description of G then there exists $\mathbf{x} \in \mathbf{A}$ such that $\mathbf{x} \in \mathbf{B}_{\phi+1} \cap \mathbf{W}_\phi$ and $\mathbf{x}\rho = x$.*

Proof. Translate corollary 2 according to the metatheorem of § 2.

Corollary 4. *Let G be a group, nilpotent of class c and $x \in \gamma_c(G)$. If $\rho: \mathbf{A} \to G$ is any description of G then there exists $\mathbf{x} \in \mathbf{A}$ such that $\mathbf{x} \in \mathbf{B}^W_{(c+1)} \cap \mathbf{N}_c^-$ and $\mathbf{x}\rho = x$.*

Proof is similar to that for corollary 1.

9. *The general process*

Definition 9·1. (A) *Let $\mathbf{x}, \mathbf{y} \in \mathbf{A}$. Then write $e: \mathbf{x} \to \mathbf{y}$ if $\mathbf{x}$ and $\mathbf{y}$ are of any of the following forms:*

(i) $\mathbf{x} = \mathbf{ab}$, $\mathbf{y} = \mathbf{ba}[\mathbf{a}, \mathbf{b}]$.
(ii) $\mathbf{x} = (\mathbf{ab})^{-1}$, $\mathbf{y} = \mathbf{b}^{-1}\mathbf{a}^{-1}$.
(iii) $\mathbf{x} = (\mathbf{a}^{-1})^{-1}$, $\mathbf{y} = \mathbf{a}$.
(iv) $\mathbf{x} = \mathbf{a}^{-1}\mathbf{a}$ *or* $\mathbf{aa}^{-1}$, $\mathbf{y} = \mathbf{1}$.
(v) $\mathbf{x} = \mathbf{a1}$ *or* $\mathbf{1a}$, $\mathbf{y} = \mathbf{a}$.
(vi) $\mathbf{x} = \mathbf{1}^{-1}$, $\mathbf{y} = \mathbf{1}$.

(vii) $\mathbf{x} = [\mathbf{a}, \mathbf{a}]$, $\mathbf{y} = \mathbf{1}$.

(viii) $\mathbf{x} = [\mathbf{a}, \mathbf{1}]$ *or* $[\mathbf{1}, \mathbf{a}]$, $\mathbf{y} = \mathbf{1}$.

(ix) $\mathbf{x} = [\mathbf{a}^{-1}, \mathbf{b}]$, $\mathbf{y} = [\mathbf{a}, \mathbf{b}]^{-1}[\mathbf{b}, \mathbf{a}, \mathbf{a}^{-1}]$ *or*
$\mathbf{x} = [\mathbf{a}, \mathbf{b}^{-1}]$, $\mathbf{y} = [\mathbf{a}, \mathbf{b}]^{-1}[\mathbf{b}, \mathbf{a}, \mathbf{b}^{-1}]$.

(x) $\mathbf{x} = [\mathbf{ab}, \mathbf{c}]$, $\mathbf{y} = [\mathbf{a}, \mathbf{c}][\mathbf{a}, \mathbf{c}, \mathbf{b}][\mathbf{b}, \mathbf{c}]$ *or*
$\mathbf{x} = [\mathbf{a}, \mathbf{bc}]$, $\mathbf{y} = [\mathbf{a}, \mathbf{c}][\mathbf{a}, \mathbf{b}][\mathbf{a}, \mathbf{b}, \mathbf{c}]$.

(xi) $\mathbf{x} = [\mathbf{a}, \mathbf{b}]$, $\mathbf{y} = [\mathbf{b}, \mathbf{a}]^{-1}$ *provided* $\mathbf{a}$ *and* $\mathbf{b}$ *are commutators and* $\mathbf{a} < \mathbf{b}$.

(xii) $\mathbf{x} = [\mathbf{c}, \mathbf{b}, \mathbf{a}]$, $\mathbf{y} = \mathbf{v}_1[\mathbf{b}, \mathbf{a}, \mathbf{c}]^{-1}\mathbf{v}_2[\mathbf{c}, \mathbf{a}, \mathbf{b}]\mathbf{v}_3$ *provided* $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{c}$ *are commutators and* $\mathbf{a} < \mathbf{b} < \mathbf{c}$, *where*

$$\mathbf{v}_1 = [\mathbf{a}, \mathbf{c}, [\mathbf{c}, \mathbf{b}]][\mathbf{c}, \mathbf{b}, [\mathbf{b}, \mathbf{a}, \mathbf{c}]],$$
$$\mathbf{v}_2 = [\mathbf{c}, \mathbf{b}, [\mathbf{b}, \mathbf{a}]][\mathbf{b}, \mathbf{a}, [\mathbf{a}, \mathbf{c}, \mathbf{b}]][\mathbf{a}, \mathbf{c}, [\mathbf{c}, \mathbf{a}, \mathbf{b}]^{-1}] \quad \textit{and}$$
$$\mathbf{v}_3 = [\mathbf{b}, \mathbf{a}, [\mathbf{a}, \mathbf{c}]][\mathbf{a}, \mathbf{c}, [\mathbf{c}, \mathbf{b}, \mathbf{a}]].$$

The notation is extended to larger expressions by recursion over their height:

(xiii) *If* $e\colon \mathbf{a}_1 \to \mathbf{a}_2$ *then* $e\colon \mathbf{a}_1^{-1} \to \mathbf{a}_2^{-1}$ *and for every* $\mathbf{b} \in \mathbf{A}$, $e\colon \mathbf{a}_1\mathbf{b} \to \mathbf{a}_2\mathbf{b}$, $e\colon \mathbf{ba}_1 \to \mathbf{ba}_2$, $e\colon [\mathbf{a}_1, \mathbf{b}] \to [\mathbf{a}_2, \mathbf{b}]$ *and* $e\colon [\mathbf{b}, \mathbf{a}_1] \to [\mathbf{b}, \mathbf{a}_2]$.

(B) *Write* $E\colon \mathbf{x} \to \mathbf{y}$ *if there exists a finite sequence* $(\mathbf{u}_i)_{i=0}^k$ $(k \geqslant 0)$ *of expressions such that* $\mathbf{u}_0 = \mathbf{x}$, $\mathbf{u}_k = \mathbf{y}$ *and* $e\colon \mathbf{u}_{i-1} \to \mathbf{u}_i$ $(1 \leqslant i \leqslant k)$.

Again the relation $E\colon \mathbf{x} \to \mathbf{y}$ is reflexive and transitive and Part (A) (xiii) of the definition holds just as well for E as for e.

Definition 9·2. (i) *For each* $\alpha \in W$ *and non-negative integer* n *the element* $\alpha(+1)^n$ *of* W *is defined recursively*: $\alpha(+1)^0 = \alpha$ *and, for* $n > 0$,

$$\alpha(+1)^n = \alpha(+1)^{n-1} + 1.$$

For any subset Φ *of* W, $\quad \Phi(+1)^n = \{\phi(+1)^n\colon \phi \in \Phi\}$.

(ii) *If* $\alpha, \beta \in W$ *then* β finitely dominates α *if* $\alpha(+1)^n < \beta$ *for every non-negative integer* n *and* β coarsely dominates α *if* $\alpha(+1)^n \not\geqslant \beta$ *for every non-negative integer* n. *If* Φ *and* Ψ *are two subsets of* W, Ψ dominates Φ *if* $\Phi(+1)^n \not\geqslant \Psi$ *for every non-negative integer* n.

Lemma 9·1. *If* $D\colon \mathbf{x} \to \mathbf{y}$ *then* $E\colon \mathbf{x} \to \mathbf{yu}$ *where* $\mathbf{u}$ *is possibly empty but if it exists* $\sigma(\mathbf{u}) \geqslant \sigma(\mathbf{x}) + 1$, $\hat\sigma(\mathbf{u}) \geqslant \hat\sigma(\mathbf{x}) + 1$, $\Sigma(\mathbf{u}) \geqslant \Sigma(\mathbf{x}) + \{1\}$ *and* $\operatorname{wt}(\mathbf{u}) \geqslant \operatorname{wt}(\mathbf{x}) + 1$.

Proof. Check the various parts of definition 8·1. For parts (A) (ii) to (viii) and (xi), $E\colon \mathbf{x} \to \mathbf{y}$ by the corresponding part of definition 9·1, so the lemma is true with $\mathbf{u}$ empty. For the remaining parts:

(A) (i) $\mathbf{x} = \mathbf{ab}$, $\mathbf{y} = \mathbf{ba}$. Then $E\colon \mathbf{x} \to \mathbf{yu}$ where

$$\mathbf{u} = [\mathbf{a}, \mathbf{b}] \quad \text{and} \quad \hat\sigma(\mathbf{u}) = \hat\sigma(\mathbf{a}) + \hat\sigma(\mathbf{b}) \geqslant \hat\sigma(\mathbf{x}) + 1$$

since $\hat\sigma(\mathbf{x}) = \hat\sigma(\mathbf{a}) \wedge \hat\sigma(\mathbf{b})$. The proofs that

$$\sigma(\mathbf{u}) \geqslant \sigma(\mathbf{x}) + 1, \quad \Sigma(\mathbf{u}) \geqslant \Sigma(\mathbf{x}) + \{1\} \quad \text{and} \quad \operatorname{wt}(\mathbf{u}) \geqslant \operatorname{wt}(\mathbf{x}) + 1$$

in this and the following cases are similar and will not be given.

(A) (ix) $\mathbf{x} = [\mathbf{a}^{-1}, \mathbf{b}]$, $\mathbf{y} = [\mathbf{a}, \mathbf{b}]^{-1}$. Then $E\colon \mathbf{x} \to \mathbf{yu}$ where

$$\mathbf{u} = [\mathbf{b}, \mathbf{a}, \mathbf{a}^{-1}] \quad \text{and} \quad \hat\sigma(\mathbf{u}) = \hat\sigma(\mathbf{b}) + \hat\sigma(\mathbf{a}) + \hat\sigma(\mathbf{a}) \geqslant \hat\sigma(\mathbf{b}) + \hat\sigma(\mathbf{a}) + 1 = \hat\sigma(\mathbf{x}) + 1.$$

The proof if $\mathbf{x} = [\mathbf{a}, \mathbf{b}^{-1}]$ is similar.

(A) (x) $\mathbf{x} = [\mathbf{ab}, \mathbf{c}]$, $\mathbf{y} = [\mathbf{a}, \mathbf{c}]\,[\mathbf{b}, \mathbf{c}]$. Then

$$E\colon \mathbf{x} \to [\mathbf{a}, \mathbf{c}]\,[\mathbf{a}, \mathbf{c}, \mathbf{b}]\,[\mathbf{b}, \mathbf{c}]$$
$$\to \mathbf{yu}$$

where $\mathbf{u} = [\mathbf{a}, \mathbf{c}, \mathbf{b}]\,[\mathbf{a}, \mathbf{c}, \mathbf{b}, [\mathbf{b}, \mathbf{c}]]$ by definition 9·1 (A) (x) and (i). But then

$$\hat{\sigma}(\mathbf{u}) = \hat{\sigma}([\mathbf{a}, \mathbf{c}, \mathbf{b}]) \geqslant \hat{\sigma}([\mathbf{a}, \mathbf{c}]) + 1 \geqslant \hat{\sigma}(\mathbf{x}) + 1.$$

The argument if $\mathbf{x} = [\mathbf{a}, \mathbf{bc}]$, $\mathbf{y} = [\mathbf{a}, \mathbf{c}]\,[\mathbf{a}, \mathbf{b}]$ is similar.

(A) (xii) $\mathbf{x} = [\mathbf{c}, \mathbf{b}, \mathbf{a}]$, $\mathbf{y} = [\mathbf{b}, \mathbf{a}, \mathbf{c}]^{-1}\,[\mathbf{c}, \mathbf{a}, \mathbf{b}]$ where $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are commutators and $\mathbf{a} < \mathbf{b} < \mathbf{c}$. Then, with the notation of definition 9·1 (A) (xii),

$$E\colon \mathbf{x} \to \mathbf{v}_1[\mathbf{b}, \mathbf{a}, \mathbf{c}]^{-1}\,\mathbf{v}_2[\mathbf{c}, \mathbf{a}, \mathbf{b}]\,\mathbf{v}_3$$
$$\to \mathbf{yu},$$

where $\mathbf{u} = \mathbf{v}_1[\mathbf{v}_1, \mathbf{y}]\,\mathbf{v}_2[\mathbf{v}_2, [\mathbf{c}, \mathbf{a}, \mathbf{b}]]\,\mathbf{v}_3$ by several applications of definition 9·1 (A) (i). To show that $\hat{\sigma}(\mathbf{u}) \geqslant \hat{\sigma}(\mathbf{x}) + 1$ it is sufficient to show that the shapes of the seven commutators which are factors of $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are all $\geqslant \hat{\sigma}(\mathbf{x}) + 1$.

(α) $\sigma([\mathbf{c}, \mathbf{b}]) \geqslant \sigma(\mathbf{c}) \geqslant \sigma(\mathbf{a})$ and so

$$\begin{aligned}\hat{\sigma}([\mathbf{a}, \mathbf{c}, [\mathbf{c}, \mathbf{b}]]) &= \hat{\sigma}([\mathbf{c}, \mathbf{a}, [\mathbf{c}, \mathbf{b}]])\\ &\geqslant \hat{\sigma}([\mathbf{c}, \mathbf{b}, \mathbf{a}, \mathbf{c}])\\ &\geqslant \hat{\sigma}([\mathbf{c}, \mathbf{b}, \mathbf{a}]) + 1.\end{aligned}$$

Here the first step follows from commutativity of addition in W and the second by definition 2·1 (v).

(β)
$$\begin{aligned}\hat{\sigma}([\mathbf{c}, \mathbf{b}, [\mathbf{b}, \mathbf{a}, \mathbf{c}]]) &\geqslant \hat{\sigma}([\mathbf{b}, \mathbf{a}, \mathbf{c}]) + 1\\ &\geqslant \hat{\sigma}([\mathbf{c}, \mathbf{b}, \mathbf{a})] + 1.\end{aligned}$$

(γ) $\sigma([\mathbf{c}, \mathbf{b}]) \geqslant \sigma(\mathbf{b}) \geqslant \sigma(\mathbf{a})$ and so

$$\begin{aligned}\hat{\sigma}([\mathbf{c}, \mathbf{b}, [\mathbf{b}, \mathbf{a}]]) &= \hat{\sigma}([\mathbf{b}, \mathbf{a}, [\mathbf{c}, \mathbf{b}]])\\ &\geqslant \hat{\sigma}([\mathbf{c}, \mathbf{b}, \mathbf{a}, \mathbf{b}])\\ &\geqslant \hat{\sigma}([\mathbf{c}, \mathbf{b}, \mathbf{a}]) + 1.\end{aligned}$$

(δ)
$$\begin{aligned}\hat{\sigma}([\mathbf{b}, \mathbf{a}, [\mathbf{a}, \mathbf{c}, \mathbf{b}]]) &\geqslant \hat{\sigma}([\mathbf{a}, \mathbf{c}, \mathbf{b}]) + 1\\ &= \hat{\sigma}([\mathbf{c}, \mathbf{a}, \mathbf{b}]) + 1\\ &= \hat{\sigma}([\mathbf{c}, \mathbf{b}, \mathbf{a}]) + 1.\end{aligned}$$

(ϵ)
$$\begin{aligned}\hat{\sigma}([\mathbf{a}, \mathbf{c}, [\mathbf{c}, \mathbf{a}, \mathbf{b}]^{-1}]) &\geqslant \hat{\sigma}([\mathbf{c}, \mathbf{a}, \mathbf{b}]) + 1\\ &= \hat{\sigma}([\mathbf{c}, \mathbf{b}, \mathbf{a}]) + 1.\end{aligned}$$

(ζ) $\sigma([\mathbf{c}, \mathbf{a}]) \geqslant \sigma(\mathbf{b}) \geqslant \sigma(\mathbf{a})$ and so

$$\begin{aligned}\hat{\sigma}([\mathbf{b}, \mathbf{a}, [\mathbf{c}, \mathbf{a}]]) &\geqslant \hat{\sigma}([\mathbf{c}, \mathbf{a}, \mathbf{b}, \mathbf{a}])\\ &\geqslant \hat{\sigma}([\mathbf{c}, \mathbf{a}, \mathbf{b}]) + 1\\ &\geqslant \hat{\sigma}([\mathbf{c}, \mathbf{b}, \mathbf{a}]) + 1.\end{aligned}$$

(η) $\hat{\sigma}([\mathbf{a}, \mathbf{c}, [\mathbf{c}, \mathbf{b}, \mathbf{a}]]) \geqslant \hat{\sigma}([\mathbf{c}, \mathbf{b}, \mathbf{a}]) + 1$ immediately.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

(A) (xiii) The result follows in this case by an easy induction over the height of $\mathbf{x}$.

(B) The result follows in this case by an easy induction over the length k of the sequence using lemma 8·5.

LEMMA 9·2. *Suppose* $E: \mathbf{x} \to \mathbf{y}$. *Then*

(i) *If* $\rho: \mathbf{A} \to G$ *is any description of a group* G *then* $\mathbf{x}\rho = \mathbf{y}\rho$.

(ii) $\sigma(\mathbf{x}) \leqslant \sigma(\mathbf{y})$, $\hat{\sigma}(\mathbf{x}) \preccurlyeq \hat{\sigma}(\mathbf{y})$, $\Sigma(\mathbf{x}) \preccurlyeq \Sigma(\mathbf{y})$ *and* $\mathrm{wt}(\mathbf{x}) \leqslant \mathrm{wt}(\mathbf{y})$.

Proof. (i) This follows easily from definition 9·1 by checking its various parts. All those parts correspond to well-known group laws, except perhaps for (A) (xii) which can be checked by expanding into a product of $\mathbf{a}\rho$, $\mathbf{b}\rho$ and $\mathbf{c}\rho$ and their inverses and cancelling.

(ii) is a corollary of lemmas 8·5 and 9·1.

LEMMA 9·3

(i) *Let* $\mathbf{x} \in \mathbf{B}_\beta$, $\mathbf{y} \in \mathbf{B}$ *and* $\sigma(\mathbf{y})+1 \geqslant \beta$. *Then* $E: \mathbf{xy} \to \mathbf{zu}$ (*either* $\mathbf{z}$ *or* $\mathbf{u}$ *possibly being empty*) *where* $\mathbf{z} \in \mathbf{B}_\beta$ *and* $\sigma(\mathbf{u}) \geqslant \beta$ *if they exist.*

(ii) *Let* $\mathbf{x} \in \hat{\mathbf{B}}_\beta$, $\mathbf{y} \in \mathbf{B}$ *and* $\hat{\sigma}(\mathbf{y})+1 \succcurlyeq \beta$. *Then* $E: \mathbf{xy} \to \mathbf{zu}$ (*either* $\mathbf{z}$ *or* $\mathbf{u}$ *possibly being empty*) *where* $\mathbf{z} \in \hat{\mathbf{B}}_\beta$ *and* $\hat{\sigma}(\mathbf{u}) \succcurlyeq \beta$ *if they exist.*

(iii) *Let* $\mathbf{x} \in \hat{\mathbf{B}}_\Psi$, $\mathbf{y} \in \mathbf{B}$ *and* $\Sigma(\mathbf{y})+\{1\} \succcurlyeq \Psi$. *Then* $E: \mathbf{xy} \to \mathbf{zu}$ (*either* $\mathbf{z}$ *or* $\mathbf{u}$ *possibly being empty*) *where* $\mathbf{z} \in \hat{\mathbf{B}}_\Psi$ *and* $\Sigma(\mathbf{u}) \succcurlyeq \Psi$ *if they exist.*

(iv) *Let* $\mathbf{x} \in \mathbf{B}_{(c)}$, $\mathbf{y} \in \mathbf{B}$ *and* $\mathrm{wt}(\mathbf{y}) \geqslant c$. *Then* $E: \mathbf{xy} \to \mathbf{zu}$ (*either* $\mathbf{z}$ *or* $\mathbf{u}$ *possibly being empty*) *where* $\mathbf{z} \in \mathbf{B}_{(c)}$ *and* $\mathrm{wt}(\mathbf{u}) \geqslant c+1$ *if they exist.*

Proof. These results are proved by an easy induction over the length (as a product) of $\mathbf{y}$ using definition 8·1 (A) (i) and lemma 9·1.

LEMMA 9·4

(i) *Suppose* $\alpha, \beta \in W$, β *does not finely dominate* α *and* $\sigma(\mathbf{x}) \geqslant \alpha$. *Then* $E: \mathbf{x} \to \mathbf{yu}$ (*either* $\mathbf{y}$ *or* $\mathbf{u}$ *possibly being empty*) *where* $\mathbf{y} \in \mathbf{B}_\beta$ *and* $\sigma(\mathbf{u}) \geqslant \beta$ *if they exist.*

(ii) *Suppose* $\alpha, \beta \in W$, β *does not coarsely dominate* α *and* $\hat{\sigma}(\mathbf{x}) \succcurlyeq \alpha$. *Then* $E: \mathbf{x} \to \mathbf{yu}$ (*either* $\mathbf{y}$ *or* $\mathbf{u}$ *possibly being empty*) *where* $\mathbf{y} \in \hat{\mathbf{B}}_\beta$ *and* $\hat{\sigma}(\mathbf{u}) \succcurlyeq \beta$ *if they exist.*

(iii) *Suppose* Φ *and* Ψ *are subsets of* W, Ψ *does not dominate* Φ *and* $\Sigma(\mathbf{x}) \succcurlyeq \Phi$. *Then* $E: \mathbf{x} \to \mathbf{yu}$ (*either* $\mathbf{y}$ *or* $\mathbf{u}$ *possibly being empty*) *where* $\mathbf{y} \in \hat{\mathbf{B}}_\Psi$ *and* $\Sigma(\mathbf{u}) \succcurlyeq \Psi$ *if they exist.*

(iv) *Let* $\mathbf{x}$ *be any expression and* c *be a positive integer. Then* $E: \mathbf{x} \to \mathbf{yu}$ (*either* $\mathbf{y}$ *or* $\mathbf{u}$ *possibly being empty*) *where* $\mathbf{y} \in \mathbf{B}_{(c)}$ *and* $\mathrm{wt}(\mathbf{u}) \geqslant c+1$ *if they exist.*

Proof will only be given for part (ii): proofs of the other three parts are similar.

Suppose n is the least non-negative integer that $\alpha(+1)^n \succcurlyeq \beta$. The argument proceeds by induction over n. If $n = 0$ then $\alpha = a(+1)^0 \succcurlyeq \beta$ so that the lemma is true with $\mathbf{y}$ empty and $\mathbf{u} = \mathbf{x}$.

Now suppose $n > 0$ and the result is true for smaller integers. Write $\beta' = \alpha(+1)^{n-1}$. Then by the inductive hypothesis, $E: \mathbf{x} \to \mathbf{y}'\mathbf{u}'$ (either $\mathbf{y}'$ or $\mathbf{u}'$ possibly being empty) where $\mathbf{y}' \in \hat{\mathbf{B}}_{\beta'}$ (and hence $\mathbf{y}' \in \hat{\mathbf{B}}_\beta$) and $\hat{\sigma}(\mathbf{u}') \succcurlyeq \beta'$ if they exist. If $\mathbf{u}'$ is empty the lemma is true with $\mathbf{y} = \mathbf{y}'$ and $\mathbf{u}$ empty. If $\mathbf{u}'$ is not empty then, by theorem 8·1, $D: \mathbf{u}' \to \mathbf{v}$ where $\mathbf{v} \in \mathbf{B}$ and $\hat{\sigma}(\mathbf{v}) \succcurlyeq \hat{\sigma}(\mathbf{u}') \succcurlyeq \beta'$. By lemma 9·1 then $E: \mathbf{u}' \to \mathbf{vz}$ ($\mathbf{z}$ possibly being empty) where $\hat{\sigma}(\mathbf{z}) \succcurlyeq \hat{\sigma}(\mathbf{u}')+1 \succcurlyeq \beta'+1 \succcurlyeq \beta$ if it exists. Thus $E: \mathbf{x} \to \mathbf{y}'\mathbf{vz}$. If $\mathbf{v}$ is empty or $\hat{\sigma}(\mathbf{v}) \succcurlyeq \beta$ the result is true with $\mathbf{y} = \mathbf{y}'$ and $\mathbf{u} = \mathbf{vz}$. Otherwise $\mathbf{y}' \in \hat{\mathbf{B}}_\beta$ and $\hat{\sigma}(\mathbf{v})+1 \succcurlyeq \beta'+1 \succcurlyeq \beta$ so, by lemma 9·3, $E: \mathbf{y}'\mathbf{v} \to \mathbf{yw}$ (either $\mathbf{y}$ or $\mathbf{w}$ possibly being empty) where $\mathbf{y} \in \hat{\mathbf{B}}_\beta$ and $\hat{\sigma}(\mathbf{w}) \succcurlyeq \beta$ if they exist: the lemma is then true with $\mathbf{u} = \mathbf{wz}$.

The reader may wonder why such pains have been taken with the exposition of this lemma since it has been patently clear for some pages that elements of a group can be described by basic expressions as the lemma implies. The reason is that the important part of the lemma is not that such expressions exist, but that they can be arrived at by only those operations listed in definition 9·1. Group theoretic results will presently be deduced by observing what can happen to certain properties of an expression under these operations.

THEOREM 9·1. THE BASIS THEOREM. (A) *Let F be an absolutely free group of rank τ on free generators $\mathcal{G} = \{g_i\}_{i<\tau}$ and let $\rho: \mathbf{A} \to F$ be the corresponding free description. Then*

(i) *For any $\alpha \in W$, $W_\alpha(F)/W_{\alpha+1}(F)$ is a free Abelian group, freely generated modulo $W_{\alpha+1}(F)$ by the set*
$$\{\mathbf{b}\rho : \mathbf{b} \textit{ is a basic commutator}, \alpha \leqslant \sigma(\mathbf{b}) < \alpha+1\}.$$

(ii) *For any $\beta \in W$, the restriction of the mapping ρ to the set $\mathbf{B}_\beta$ of basic expressions involving only commutators of fine shape $< \beta$ is one-to-one into $F/W_\beta(F)$ modulo $W_\beta(F)$.*

(iii) *Provided β does not finely dominate α, the restriction of ρ to the set $\mathbf{W}_\alpha \cap \mathbf{B}_\beta$ of basic expressions involving only commutators of fine shape $\geqslant \alpha$ and $< \beta$ is one-to-one* onto *the factor $W_\alpha(F)/W_\beta(F)$ modulo $W_\beta(F)$.*

(B) *With the same conditions as for* (A),

(i) *For any $\alpha \in W$, $\hat{W}_\alpha(F)/\hat{W}_{\alpha+1}(F)$ is a free Abelian group, freely generated modulo $\hat{W}_{\alpha+1}(F)$ by the set* $\{\mathbf{b}\rho : \mathbf{b}$ *is a basic commutator*, $\alpha \leqslant \hat{\sigma}(\mathbf{b}) \not\geqslant \alpha+1\}$.

(ii) *For any $\beta \in W$, the restriction of ρ to the set $\hat{\mathbf{B}}_\beta$ of basic expressions involving only commutators of coarse shape $\not\geqslant \beta$ is one-to-one into $F/\hat{W}_\beta(F)$ modulo $\hat{W}_\beta(F)$.*

(iii) *Provided β does not coarsely dominate α, the restriction of ρ to the set $\hat{\mathbf{W}}_\alpha \cap \hat{\mathbf{B}}_\beta$ of basic expressions involving only commutators of coarse shape $\geqslant \alpha$ and $\not\geqslant \beta$ is one-to-one* onto *the factor $\hat{W}_\alpha(F)/\hat{W}_\beta(F)$ modulo $\hat{W}_\beta(F)$.*

(C) *With the same conditions as for* (A),

(i) *For any subset Φ of W, $\hat{W}_\Phi(F)/\hat{W}_{\Phi+\{1\}}(F)$ is a free Abelian group, freely generated modulo $\hat{W}_{\Phi+\{1\}}(F)$ by the set* $\{\mathbf{b}\rho : \mathbf{b}$ *is a basic commutator*, $\Phi \leqslant \Sigma(\mathbf{b}) \not\geqslant \Phi+\{1\}\}$.

(ii) *For any $\Psi \subseteq W$, the restriction of the mapping ρ to the set $\hat{\mathbf{B}}_\Psi$ of basic expressions involving only commutators of shape set $\not\geqslant \Psi$ is one-to-one into $F/\hat{W}_\Psi(F)$ modulo $\hat{W}_\Psi(F)$.*

(iii) *Provided that Ψ does not dominate Φ, the restriction of ρ to the set $\hat{\mathbf{W}}_\Phi \cap \hat{\mathbf{B}}_\Psi$ of basic expressions involving only commutators of shape set $\geqslant \Phi$ and $\not\geqslant \Psi$ is one-to-one* onto *the factor $\hat{W}_\Phi(F)/\hat{W}_\Psi(F)$ modulo $\hat{W}_\Psi(F)$.*

(D) *Let G be a group, free with respect to being nilpotent of class c, of rank τ and freely generated by $\mathcal{G} = \{g_i\}_{i<\tau}$.*

Let $\rho: \mathbf{A} \to G$ be the corresponding free description. Then

(i) *$\gamma_c(G)$ is a free Abelian group, freely generated by the set*
$$\{\mathbf{b}\rho : \mathbf{b} \textit{ is a basic commutator}, \operatorname{wt}(\mathbf{b}) = c\}.$$

(ii) *For any $\alpha \in W$ and $\Phi \subseteq W$, the subgroups $W_\alpha(G) \cap \gamma_c(G)$, $\hat{W}_\alpha(G) \cap \gamma_c(G)$ and $\hat{W}_\Phi(G) \cap \gamma_c(G)$ are free Abelian groups, freely generated by the sets*
$$\{\mathbf{b}\rho : \mathbf{b} \textit{ is a basic commutator}, \operatorname{wt}(\mathbf{b}) = c, \sigma(\mathbf{b}) \geqslant \alpha\},$$
$$\{\mathbf{b}\rho : \mathbf{b} \textit{ is a basic commutator}, \operatorname{wt}(\mathbf{b}) = c, \sigma(\mathbf{b}) \geqslant \alpha\},$$
and $$\{\mathbf{b}\rho : \mathbf{b} \textit{ is a basic commutator}, \operatorname{wt}(\mathbf{b}) = c, \Sigma(\mathbf{b}) \geqslant \Phi\},$$
respectively.

(iii) *For any $\alpha, \beta \in W$ the restriction of ρ to the set $\mathbf{W}_\alpha \cap \mathbf{B}_\beta \cap \mathbf{B}_{(c)}$ is one-to-one onto the factor $W_\alpha(G)/W_\beta(G)$ modulo $W_\beta(G)$ and the restriction of ρ to the set $\hat{\mathbf{W}}_\alpha \cap \hat{\mathbf{B}}_\beta \cap \mathbf{B}_{(c)}$ is one-to-one onto the factor $\hat{W}_\alpha(G)/\hat{W}_\beta(G)$ modulo $\hat{W}_\beta(G)$. For any subsets Φ and Ψ of W, the restriction of ρ to the set $\hat{\mathbf{W}}_\Phi \cap \hat{\mathbf{B}}_\Psi \cap \mathbf{B}_{(c)}$ is one-to-one onto the factor $\hat{W}_\Phi(G)/\hat{W}_\Psi(G)$ modulo $\hat{W}_\Psi(G)$.*

Proof. It will be noticed that the statement of the theorem contains analogous sets of results for the fine shape, coarse shape and shape set subgroups. In each case the results for the coarse shape subgroups can be inferred from the corresponding ones for the shape set subgroups by substituting $\Phi = \{\alpha\}$ and $\Psi = \{\beta\}$ and then the results for the fine shape subgroups by invocation of the metatheorem of § 2. Thus only proofs of the results for shape set subgroups need be given.

(D) (i) It is clearly sufficient to prove this statement when τ is finite. Let $x \in \gamma_c(G)$. Then there exists $\mathbf{x} \in \mathbf{A}$ such that $\mathbf{x}\rho = x$ and wt $(\mathbf{x}) \geqslant c$. By lemma 9·4 (iv), E: $\mathbf{x} \to \mathbf{yu}$ (either $\mathbf{y}$ or $\mathbf{u}$ possibly being empty) where $\mathbf{y} \in \mathbf{B}_{(c)}$ and wt $(\mathbf{u}) \geqslant c+1$ if they exist. Then

$$\mathbf{x}\rho = (\mathbf{yu})\rho = \mathbf{y}\rho$$

by lemma 9·2 (i). If $\mathbf{y}$ is empty or $\mathbf{1}$ then $x = \mathbf{y}\rho = 1$, otherwise it is a basic expression of the form $\mathbf{y} = \mathbf{b}_1^{\alpha_1}\mathbf{b}_2^{\alpha_2}\dots\mathbf{b}_k^{\alpha_k}$. Since $\mathbf{y} \in \mathbf{B}_{(c)}$, each $\mathbf{b}_i$ is of weight $\leqslant c$; but also, by lemma 9·2 (ii), wt $(\mathbf{y}) \geqslant$ wt $(\mathbf{yu}) \geqslant$ wt $(\mathbf{x}) \geqslant c$ so each $\mathbf{b}_i$ is of weight exactly c. This proves that $\gamma_c(G)$ is generated by $\mathbf{X}\rho$ where $\mathbf{X}$ is the set of basic commutators of weight exactly c. By the conventional theory ρ maps the finite set $\mathbf{Y}$ of N^--basic commutators one-to-one into $\gamma_c(G)$ which is free Abelian, freely generated by $\mathbf{Y}\rho$. By theorem 7·1, $|\mathbf{X}| = |\mathbf{Y}|$; the result follows.

(D) (ii) Now write $\mathbf{X}$ for the set of basic commutators of weight exactly c and shape set $\geqslant \Phi$. By virtue of part (D) (i), it is sufficient to prove that $\mathbf{X}\rho$ generates $\hat{W}_\Phi(G) \cap \gamma_c(G)$. First, $\mathbf{X}\rho \subseteq \hat{W}_\Phi(G) \cap \gamma_c(G)$ by definition 3·2. Now suppose $x \in \hat{W}_\Phi(G) \cap \gamma_c(G)$. Then since $x \in \hat{W}_\Phi(G)$ there exists $\mathbf{x}' \in \mathbf{A}$ such that $\mathbf{x}'\rho = x$ and $\Sigma(\mathbf{x}) \geqslant \Phi$. By lemma 9·4 (iv), E: $\mathbf{x}' \to \mathbf{xu}$ (either $\mathbf{x}$ or $\mathbf{u}$ possibly being empty) where $\mathbf{x} \in \mathbf{B}_{(c)}$ and wt $(\mathbf{u}) \geqslant c+1$ if they exist. If $\mathbf{x}$ is empty or $\mathbf{1}$ then $x = (\mathbf{xu})\rho = 1$. Otherwise $x = \mathbf{x}\rho$ and $\mathbf{x}$ is a basic expression of the form $\mathbf{x} = \mathbf{b}_1{}^{\alpha_1}\mathbf{b}_2{}^{\alpha_2}\dots\mathbf{b}_k{}^{\alpha_k}$. By lemma 9·2, $\Sigma(\mathbf{x}) \geqslant \Sigma(\mathbf{xu}) \geqslant \Sigma(\mathbf{x}') \geqslant \Phi$ so each $\mathbf{b}_i$ has shape set $\geqslant \Phi$. Since $\mathbf{x} \in \mathbf{B}_{(c)}$ each $\mathbf{b}_i$ is of weight $\leqslant c$. Finally suppose the $\mathbf{b}_i$ are not all of weight exactly c. Then let c' be the least weight of the $\mathbf{b}_i$, so that $c' < c$. Form the expression $\mathbf{x}''$ by deleting all commutators of weight $> c'$ from $\mathbf{x}$. Then $\mathbf{x}''\rho = x$ modulo $\gamma_{c'+1}(G)$ and $\mathbf{x}''$ is a non-trivial basic expression involving only commutators of weight exactly c' so, by (D) (i), $x \notin \gamma_{c'+1}(G)$ which contradicts $x \in \gamma_c(G)$.

(D) (iii) By induction over c: for $c = 0$ this is trivial. Now suppose $c > 0$ and the statement is true for smaller c. By definition 3·2, ρ maps $\hat{\mathbf{W}}_\Phi \cap \hat{\mathbf{B}}_\Psi \cap \mathbf{B}_{(c)}$ into $\hat{W}_\Phi(G)$. Now suppose $x \in G$. Then by the inductive hypothesis, there exists $\mathbf{x}' \in \hat{\mathbf{W}}_\Phi \cap \hat{\mathbf{B}}_\Psi \cap \mathbf{B}_{(c-1)}$ such that $\mathbf{x}'\rho = x$ modulo $\gamma_c(G)$. $\hat{W}_\Psi(g)$. Thus there exists $u \in \gamma_c(G)$ such that $x = (\mathbf{x}'\rho)u$. But x and $\mathbf{x}'\rho \in \hat{W}_\Phi(G)$ so $u \in \gamma_c(G) \cap \hat{W}_\Phi(G)$.

By (D) (ii) there exists $\mathbf{u} \in \mathbf{B}_{(c)} \cap \hat{\mathbf{W}}_\Phi$ of weight $\geqslant c$ such that $\mathbf{u}\rho = u$. Then $(\mathbf{x}'\mathbf{u})\rho = x$ and by lemma 9·3, E: $\mathbf{x}'\mathbf{u} \to \mathbf{xz}$ (either $\mathbf{x}$ or $\mathbf{z}$ possibly being empty) where $\mathbf{x} \in \mathbf{B}_{(c)}$ and wt $(\mathbf{z}) \geqslant c+1$ if they exist. If $\mathbf{x}$ is empty or $\Sigma(\mathbf{x}) \geqslant \Psi$ then $x = 1 = \mathbf{1}\rho$ modulo $\hat{W}_\Psi(G)$ and $\mathbf{1} \in \hat{\mathbf{W}}_\Phi \cap \hat{\mathbf{B}}_\Psi \cap \mathbf{B}_{(c)}$. Otherwise $x = \mathbf{x}\rho$ and $\mathbf{x} \in \mathbf{B}_{(c)}$. But $\Sigma(\mathbf{x}) \geqslant \Sigma(\mathbf{xz}) \geqslant \Sigma(\mathbf{x}'\mathbf{u}) \geqslant \Phi$ by lemma 9·2 (ii) so $\mathbf{x} \in \hat{\mathbf{W}}_\Phi$ and deleting any commutators of shape set $\geqslant \Psi$ in $\mathbf{x}$ will not change $\mathbf{x}\rho$ modulo $\hat{W}_\Psi(G)$, so $\mathbf{x}$ may be chosen $\in \hat{\mathbf{B}}_\Psi$. This proves that ρ maps $\hat{\mathbf{W}}_\Phi \cap \hat{\mathbf{B}}_\Psi \cap \mathbf{B}_{(c)}$ onto

$\hat{W}_\Phi(G)/\hat{W}_\Psi(G)$ modulo $\hat{W}_\Psi(G)$. Now suppose $\mathbf{x}$ and $\mathbf{y} \in \hat{\mathbf{W}}_\Phi \cap \hat{\mathbf{B}}_\Psi \cap \mathbf{B}_{(c)}$ and $\mathbf{x}\rho = \mathbf{y}\rho$ modulo $\hat{W}_\Psi(G)$. From $\mathbf{x}$ form the expressions $\mathbf{x}_1$ by deleting all commutators of weight c and $\mathbf{x}_2$ by deleting all commutators of weight $< c$. From $\mathbf{y}$ form $\mathbf{y}_1$ and $\mathbf{y}_2$ similarly. Then

$$\mathbf{x}_1, \mathbf{y}_1 \in \hat{\mathbf{W}}_\Phi \cap \hat{\mathbf{B}}_\Psi \cap \mathbf{B}_{(c-1)} \quad \text{and} \quad \mathbf{x}_1\rho = \mathbf{y}_1\rho \text{ modulo } \gamma_c(G).$$

$\hat{W}_\Psi(G)$ so by the inductive hypothesis $\mathbf{x}_1 = \mathbf{y}_1$. But x can be converted into $\mathbf{x}_1\mathbf{x}_2$ by interchanging the position of commutators of weight exactly c with others, so $(\mathbf{x}_1\mathbf{x}_2)\rho = \mathbf{x}\rho$. Similarly, $(\mathbf{y}_1\mathbf{y}_2)\rho = \mathbf{y}\rho$ and thus $\mathbf{x}_2\rho = \mathbf{y}_2\rho$ modulo $\hat{W}_\Psi(G)$. But $\mathbf{x}_2$ and $\mathbf{y}_2$ are basic expressions involving only commutators of weight exactly c and shape set $\not\geqslant \Psi$. Thus $\mathbf{x}_2 = \mathbf{y}_2$ by (D) (i). Referring to definition 6·2 (B), $\mathbf{x} = \mathbf{y}$.

(C) (ii) Let $\mathbf{x}, \mathbf{y} \in \hat{\mathbf{B}}_\Psi$ and $\mathbf{x}\rho = \mathbf{y}\rho$ modulo $\hat{W}_\Psi(F)$. Since $\mathbf{x}$ and $\mathbf{y}$ are products of finite length, there exists an integer c such that $\mathbf{x}, \mathbf{y} \in \mathbf{B}_{(c)}$ and, since $\hat{\mathbf{B}}_{\{1\}} = \mathbf{A}$, $\mathbf{x}, \mathbf{y} \in \hat{\mathbf{W}}_{\{1\}}$ trivially. Thus $\mathbf{x}, \mathbf{y} \in \hat{\mathbf{W}}_{\{1\}} \cap \hat{\mathbf{B}}_\Psi \cap \mathbf{B}_{(c)}$ and $\mathbf{x}\rho = \mathbf{y}\rho$ modulo $\hat{W}_\Psi(F).\ \gamma_c(F)$. Hence $\mathbf{x} = \mathbf{y}$ by (D) (iii).

(C) (iii) By virtue of (C) (ii) it is sufficient to show that the restriction of ρ to $\hat{\mathbf{W}}_\Phi \cap \hat{\mathbf{B}}_\Psi$ is onto $\hat{W}_\Phi(F)$ modulo $\hat{W}_\Psi(F)$. Suppose then that $x \in \hat{W}_\Phi(F)$. Then there exists an expression $\mathbf{x}$ such that $\mathbf{x}\rho = x$ and $\Sigma(\mathbf{x}) \geqslant \Phi$. Then by lemma 9·4, E: $\mathbf{x} \to \mathbf{yu}$ (either $\mathbf{y}$ or $\mathbf{u}$ possibly being empty) where $\mathbf{y} \in \hat{\mathbf{B}}_\Phi$ and $\Sigma(\mathbf{u}) \geqslant \Psi$ if they exist. If $\mathbf{y}$ is empty or $\mathbf{1}$, $x = \mathbf{u}\rho = 1$ modulo $\hat{W}_\Psi(F)$. Otherwise $x = \mathbf{y}\rho$ and $\mathbf{y} \in \hat{\mathbf{W}}_\Phi \cap \hat{\mathbf{B}}_\Psi$ by lemma 9·2.

(C) (i) That $\hat{W}_\Phi(F)/\hat{W}_{\Phi+\{1\}}(F)$ is Abelian is obvious. The remainder of this part is now a corollary of (C) (ii) and (iii).

10. *Lie rings*

The theory developed in this paper may be applied to Lie rings in a simplified form. This will not be expounded in detail here; instead the two crucial lemmas will be given and then the details may safely be left to the reader.

In the language used in this paper, a Lie ring may be defined conveniently as follows.

Definition 10·1. *A Lie ring L is a describable algebra in which the effects of the operators ϵ, ν, μ and χ are written*

$$\epsilon = 0$$

$$\left.\begin{aligned} x\nu &= -x \\ xy\mu &= x+y \\ xy\chi &= xy \end{aligned}\right\} \text{ for any } \quad x, y \in L$$

and with the laws:

(i) *L is an Abelian group with respect to μ, ν and ϵ,*
(ii) $x(y+z) = xy+xz$ *and* $(x+y)z = xz+yz$,
(iii) $xx = 0$,
(iv) $(xy)z+(yz)x+(zx)y = 0$.

Immediate consequences of these laws are

(v) $xy = -yx$,
(vi) $x(-y) = (-x)y = -(xy)$,
(vii) $x0 = 0x = 0$.

Hence the special collecting process operates anywhere in a Lie ring, in the following sense.

LEMMA 10·1. *Let L be a Lie ring and* $\rho: \mathbf{A} \to L$ *a description of L. If* $D: \mathbf{x} \to \mathbf{y}$ *then* $\mathbf{x}\rho = \mathbf{y}\rho$.

Since a Lie ring is defined in terms of laws, the idea of a *free Lie ring* is tenable. M. Hall Jr. (1950) has proved (restating his theorem 3·1 in the language of this paper):

If L is a free Lie ring, freely generated by the set $\mathscr{G} = \{g_i\}_{i<\tau}$, *and* $\rho: \mathbf{A} \to L$ *is the corresponding free description of L, then the* N^-*-basic commutators are mapped one-to-one into L by* ρ *and their images form a basis for L (that is, they generate L freely* qua *free Abelian group).*

This makes possible

LEMMA 10·2. *Let L be a free Lie ring,* $\rho: \mathbf{A} \to L$ *a free description of L and W any shape range. Then the W-basic commutators are mapped one-to-one into L by* ρ *and their images form a basis for L*

Proof. (i) Let $x \in L$. Then there exists $\mathbf{x} \in \mathbf{A}$ such that $\mathbf{x}\rho = x$ and then, by theorem 8·1, $D: \mathbf{x} \to \mathbf{y}$ where $\mathbf{y}$ is a basic expression. By lemma 10·1, $x = \mathbf{y}\rho$. If $\mathbf{y} = \mathbf{1}$ then $x = 0$; otherwise $\mathbf{y}$ is of the form $\mathbf{y} = \mathbf{b}_1^{\alpha_1}\mathbf{b}_2^{\alpha_2}\ldots\mathbf{b}_k^{\alpha_k}$, where each $\mathbf{b}_i$ is a basic commutator, and then $x = \alpha_1(\mathbf{b}_1\rho) + \alpha_2(\mathbf{b}_2\rho) + \ldots + \alpha_k(\mathbf{b}_k\rho)$. This proves that the images of basic commutators under ρ generate L *qua* Abelian group.

(ii) By the same argument, using part (iii) of theorem 8·1 also, the set

$$\{\mathbf{b}\rho: \mathbf{b} \text{ is a basic commutator, wt}(\mathbf{b}) = c\}$$

generates $\gamma_c(L)/\gamma_{c+1}(L)$ modulo $\gamma_{c+1}(L)$ and then, by the theorem of M. Hall Jr. quoted above and theorem 7·1, this set is a basis for $\gamma_c(L)/\gamma_{c+1}(L)$ modulo $\gamma_{c+1}(L)$.

(iii) It is now shown that the restriction of ρ to $\mathbf{B}_{(c)}$ maps this set one-to-one onto $L/\gamma_{c+1}(L)$ modulo $\gamma_{c+1}(L)$ by induction over c. Let $x \in L$. Then by (i) there exists a W-basic expression $\mathbf{x}$ such that $\mathbf{x}\rho = x$. If wt$(\mathbf{x}) \geqslant c+1$ then $x = 0 = \mathbf{1}\rho$ modulo $\gamma_{c+1}(L)$ and $\mathbf{1} \in \mathbf{B}_{(c)}$. Otherwise form the expression $\mathbf{x}'$ by deleting all commutators of weight $\geqslant c+1$ from $\mathbf{x}$. Then $\mathbf{x}' \in \mathbf{B}_{(c)}$ and $\mathbf{x}'\rho = x$ modulo $\gamma_{c+1}(L)$. Thus ρ maps $\mathbf{B}_{(c)}$ onto L modulo $\gamma_{c+1}(L)$. Now suppose $\mathbf{x}, \mathbf{y} \in \mathbf{B}_{(c)}$ and $\mathbf{x}\rho = \mathbf{y}\rho$ modulo $\gamma_{c+1}(L)$. From $\mathbf{x}$ form the expressions $\mathbf{x}_1$ by deleting all commutators of weight c and $\mathbf{x}_2$ by deleting all commutators of weight $< c$. From $\mathbf{y}$ form $\mathbf{y}_1$ and $\mathbf{y}_2$ in the same manner. Then $\mathbf{x}_1, \mathbf{y}_1 \in \mathbf{B}_{(c-1)}$ and $\mathbf{x}_1\rho = \mathbf{y}_1\rho$ modulo $\gamma_c(L)$ so by the inductive hypothesis $\mathbf{x}_1 = \mathbf{y}_1$. But $(\mathbf{x}_1\mathbf{x}_2)\rho = \mathbf{x}\rho = \mathbf{y}\rho = (\mathbf{y}_1\mathbf{y}_2)\rho$ modulo $\gamma_{c+1}(L)$ so $\mathbf{x}_2\rho = \mathbf{y}_2\rho$ modulo $\gamma_{c+1}(L)$. But $\mathbf{x}_2$ and $\mathbf{y}_2$ are basic expressions involving only commutators of weight exactly c, so $\mathbf{x}_2 = \mathbf{y}_2$ by part (ii) of this proof. Now $\mathbf{x}_1 = \mathbf{y}_1$ and $\mathbf{x}_2 = \mathbf{y}_2$ so $\mathbf{x} = \mathbf{y}$.

(iv) Finally, let $\mathbf{x}$ and $\mathbf{y}$ be two basic expressions such that $\mathbf{x}\rho = \mathbf{y}\rho$. Then there exists an integer c such that $\mathbf{x}, \mathbf{y} \in \mathbf{B}_{(c)}$; then $\mathbf{x}\rho = \mathbf{y}\rho$ modulo $\gamma_{c+1}(L)$ so $\mathbf{x} = \mathbf{y}$ by part (iii) of this proof.

11. *Fine shape subgroups as products of coarse ones*

It has been mentioned that, for any expression $\mathbf{x}$, $\hat{\sigma}(\mathbf{x}) \leqslant \sigma(\mathbf{x})$. It is not difficult to construct examples for which this order relation is strict. On the other hand, a weakened form of the opposite relation may be established as follows:

LEMMA 11·1. *Let* $\mathbf{x} \in \mathbf{A}$. *Then there exist* $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k \in \mathbf{A}$ *such that* $E: \mathbf{x} \to \mathbf{u}_1\mathbf{u}_2\ldots\mathbf{u}_k$ *and* $\hat{\sigma}(\mathbf{u}_i) \geqslant \sigma(\mathbf{x})$ $(1 \leqslant i \leqslant k)$.

Proof. By induction over the height of $\mathbf{x}$.

LEMMA 11·2. *Let W be a shape range and* $\alpha \in W$. *Then, for any group G,*

$$W_\alpha(G) = \prod_{\xi \geqslant \alpha} \hat{W}_\xi(G).$$

Proof. For any $\xi \geqslant \alpha$, $\hat{W}_\xi(G) \leqslant W_\xi(G) \leqslant W_\alpha(G)$ by theorem 3·1 (iv) and (vii). Thus

$$\prod_{\xi \geqslant \alpha} \hat{W}_\xi(G) \leqslant W_\alpha(G).$$

Now suppose $x \in W_\alpha(G)$. Then there exists $\mathbf{x} \in \mathbf{A}$ such that $\mathbf{x}\rho = x$ and $\sigma(\mathbf{x}) \geqslant \alpha$. By lemma 12·1, E: $\mathbf{x} \to \mathbf{u}_1 \mathbf{u}_2 \ldots \mathbf{u}_k$ where $\hat{\sigma}(\mathbf{u}_i) \geqslant \sigma(\mathbf{x}) \geqslant \alpha$ $(1 \leqslant i \leqslant k)$. Then, writing $\hat{\sigma}(\mathbf{u}_i) = \xi_i$,

$$x = \mathbf{x}\rho = \mathbf{u}_1\rho \,.\, \mathbf{u}_2\rho \ldots \mathbf{u}_k\rho \in \prod_{\xi \geqslant \alpha} \hat{W}_\xi(G).$$

This lemma, together with theorem 3·1 (viii), shows that the three types of shape subgroups can all be represented as shape-set subgroups.

12. *Distinctness of shape subgroups*

The question as to when the various shape subgroups are nontrivial and when they are different from one another for an arbitrary group G is obviously a very complicated one, and one which depends very much on the special properties of G. However, if $G \cong F_\tau$ is an absolutely free group of rank at least 3 the answer is very simple: for any two subsets Φ and Ψ of a shape range W, $\hat{W}_\Phi(G) = \hat{W}_\Psi(G)$ if and only if Φ and Ψ are equivalent under the pre-order $\leqslant$. This, together with the results of the previous section, allows all the shape subgroups to be compared; in particular, for $\alpha, \beta \in W$ the propositions

$$\alpha = \beta, \quad \hat{W}_\alpha(G) = \hat{W}_\beta(G) \quad \text{and} \quad W_\alpha(G) = W_\beta(G)$$

are equivalent. This is proved in this section.

The proof when the number of generators is 3 depends on the property of a shape range W that any $\alpha \in W$ other than 1 or ∞ may be written in the form

$$\alpha = 1 + \alpha_0 + \alpha_1 + \ldots + \alpha_k,$$

where $k \geqslant 0$, $1 = \alpha_0 \leqslant \alpha_1 \leqslant \ldots \leqslant \alpha_k$ and for each r $(0 \leqslant r < k)$, $1 + \alpha_0 + \alpha_1 + \ldots + \alpha_r \geqslant \alpha_{r+1}$. The proof of this in turn involves something very much like a collecting process operating upon such 'formal sums' in W. This appalling prospect may be circumvented however by considering the properties of commutators in $\mathbf{A}$ when the number of generators is infinite.

Definition 12·1. *For any commutator* $\mathbf{c}$ *and for any* $i < \tau$, *the* number of times $\mathbf{c}$ mentions $\mathbf{g}_i$, *written* $\mu_i(\mathbf{c})$, *is defined recursively over the weight of* $\mathbf{c}$:

(i) $\mu_i(\mathbf{g}_j) = 1$ *if* $i = j$,
$\phantom{(i) \mu_i(\mathbf{g}_j)} = 0$ *if* $i \neq j$.

(ii) $\mu_i([\mathbf{a}, \mathbf{b}]) = \mu_i(\mathbf{a}) + \mu_i(\mathbf{b})$.

Lemma 12·1. *Suppose* $\mathbf{c}$ *is a non-basic commutator which mentions each generator at most once* (*that is,* $\mu_i(\mathbf{c}) = 0$ *or* 1 *for every* $i < \tau$). *Then there exists a commutator* $\mathbf{c}'$ *such that* $\mathbf{c}' < \mathbf{c}$, $\hat{\sigma}(\mathbf{c}') = \hat{\sigma}(\mathbf{c})$ *and, for each* $i < \tau$, $\mu_i(\mathbf{c}') = \mu_i(\mathbf{c})$.

Proof. The argument proceeds by induction over the weight of $\mathbf{c}$. Since $\mathbf{c}$ is non-basic, at least one of the conditions of definition 6·2 (A) must fail for this commutator.

Suppose $\mathbf{c} = [\mathbf{c}_1, \mathbf{c}_2]$ and $\mathbf{c}_1$ is non-basic. Then clearly $\mathbf{c}_1$ mentions each generator at most once so there exists $\mathbf{c}_1'$ such that $\mathbf{c}_1' < \mathbf{c}_1$, $\hat{\sigma}(\mathbf{c}_1') = \hat{\sigma}(\mathbf{c}_1)$ and, for each $i < \tau$, $\mu_i(\mathbf{c}_1') = \mu_i(\mathbf{c}_1)$. Then $\mathbf{c}' = [\mathbf{c}_1', \mathbf{c}_2]$ is the required commutator. The proof if $\mathbf{c}_2$ is non-basic is similar.

Now suppose $\mathbf{c} = [\mathbf{c}_1, \mathbf{c}_2]$ and $\mathbf{c}_1 \leqslant \mathbf{c}_2$. Then, since $\mathbf{c}$ mentions each generator at most once, $\mathbf{c}_1 \neq \mathbf{c}_2$ so $\mathbf{c}_1 < \mathbf{c}_2$. Then $\mathbf{c}' = [\mathbf{c}_2, \mathbf{c}_1]$ is the required commutator.

Finally, suppose $\mathbf{c} = [\mathbf{c}_{11}, \mathbf{c}_{12}, \mathbf{c}_2]$ and $\mathbf{c}_{12} > \mathbf{c}_2$. By virtue of the first possibility considered, it may be assumed that $\mathbf{c}_{11} > \mathbf{c}_{12}$. Then, by the proof of lemma 8·3, $\mathbf{c}' = [\mathbf{c}_{11}, \mathbf{c}_2, \mathbf{c}_{12}]$ is the required commutator.

Corollary. *Let the number of generators of* $\mathbf{A}$ *be infinite and let* $\alpha \in W$ $(\alpha \neq \infty)$. *Then there exists a basic commutator of shape exactly* α *in* A.

Proof. By an easy induction over α, there exists a commutator $\mathbf{c} \in \mathbf{A}$ which mentions each generator at most once for which $\hat{\sigma}(\mathbf{c}) = \alpha$. If $\mathbf{c}$ is non-basic it may be replaced by an earlier commutator with the same properties. But the set of commutators is well-ordered by $\leqslant$ so there exists a basic commutator with these properties.

Lemma 12·2. *If* $\alpha \in W$, $\alpha \neq 1$ *or* ∞, *then it may be written in the form*

$$\alpha = 1+\alpha_0+\alpha_1+\ldots+\alpha_k \quad (k \geqslant 0),$$

where $1 = \alpha_0 \leqslant \alpha_1 \leqslant \ldots \leqslant \alpha_k$ *and for each* r $(0 \leqslant r < k)$, $1+\alpha_0+\alpha_1+\ldots+\alpha_r \geqslant {}_{r+1}$.

Proof. Form an algebra $\mathbf{A}$ of expressions with an infinite number of generators. Then by the corollary to lemma 12·1 there exists a basic commutator $\mathbf{c}$ of shape exactly α in $\mathbf{A}$. But then $\mathbf{c}$ may be written in the form

$$\mathbf{c} = [\mathfrak{g}_i, \mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k] \quad (k \geqslant 0),$$

where $\mathfrak{g}_i > \mathbf{a}_0 \leqslant \mathbf{a}_1 \leqslant \mathbf{a}_2 \leqslant \ldots \leqslant \mathbf{a}_k$ and for each r $(0 \leqslant r < k)$,

$$[\mathfrak{g}_i, \mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_r] > \mathbf{a}_{r+1}.$$

Writing $\alpha_i = \sigma(\mathbf{a}_i)$ $(0 \leqslant i \leqslant k)$ the lemma follows.

This result having been obtained, the three-generator case may be considered.

Lemma 12·3. *Let* W *be a shape range and* α *be an element of* W *other than* ∞. *If the number* τ *of generators of* $\mathbf{A}$ *is at least* 3 *then there exist three distinct basic commutators of shape exactly* α *in* $\mathbf{A}$.

Proof. When $\alpha = 1$ the result is trivial—$\mathfrak{g}_0$, $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are the required commutators.

Now suppose that $\alpha > 1$. It is proved by induction over k that, however α may be written in the form

$$\alpha = 1+\alpha_0+\alpha_1+\ldots+\alpha_k,$$

where $\alpha_r \leqslant \alpha_{r+1} \leqslant 1+\alpha_0+\ldots+\alpha_r$ $(0 \leqslant r < k)$ and $\alpha_0 = 1$, there exist at least three basic commutators of the form

$$\mathbf{a} = [\mathfrak{g}_i, \mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_k]$$

in $\mathbf{A}$ where $\sigma(\mathbf{a}_r) = \alpha_r$ $(0 \leqslant r \leqslant k)$.

Suppose then that α may be written in the form given above. If $k = 0$ then $\alpha = 1+1$ and the required commutators are $[\mathfrak{g}_1, \mathfrak{g}_0]$, $[\mathfrak{g}_2, \mathfrak{g}_0]$ and $[\mathfrak{g}_2, \mathfrak{g}_1]$. Now suppose that $k \geqslant 1$. Then $\alpha_{k-1} \leqslant \alpha_k \leqslant 1+\alpha_0+\ldots+\alpha_{k-1}$ and there exist three basic commutators

$$\mathbf{a} = [\mathfrak{g}_i, \mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_{k-1}], \quad \mathbf{a}' = [\mathfrak{g}_{i'}, \mathbf{a}'_0, \mathbf{a}'_1, \ldots, \mathbf{a}'_{k-1}] \quad \text{and} \quad \mathbf{a}'' = [\mathfrak{g}_{i''}, \mathbf{a}''_0, \mathbf{a}''_1, \ldots, \mathbf{a}''_{k-1}],$$

where $\sigma(\mathbf{a}_i) = \sigma(\mathbf{a}'_i) = \sigma(\mathbf{a}''_i) = \alpha_i$ $(0 \leqslant r < k)$. But $\alpha_{k-1} < 1+\alpha_0+\alpha_1+\ldots+\alpha_{k-1}$ so there are three possibilities:

(i) If $\alpha_{k-1} = \alpha_k < 1+\alpha_0+\alpha_1+\ldots+\alpha_{k-1}$ then the required commutators are $[\mathbf{a}, \mathbf{a}_{k-1}]$, $[\mathbf{a}', \mathbf{a}'_{k-1}]$ and $[\mathbf{a}'', \mathbf{a}''_{k-1}]$.

(ii) If $\alpha_{k-1} < \alpha_k < 1+\alpha_0+\alpha_1+\ldots+\alpha_{k-1}$, let $\mathbf{b}$ be any one of the three basic commutators of shape α_k. Then the required commutators are $[\mathbf{a},\mathbf{b}]$, $[\mathbf{a}',\mathbf{b}]$ and $[\mathbf{a}'',\mathbf{b}]$.

(iii) Finally, if $\alpha_{k-1} < \alpha_k = 1+\alpha_0+\alpha_1+\ldots+\alpha_{k-1}$ then since $\mathbf{a}$, $\mathbf{a}'$ and $\mathbf{a}''$ are distinct it may be assumed that $\mathbf{a} < \mathbf{a}' < \mathbf{a}''$. Then the required commutators are $[\mathbf{a}',\mathbf{a}]$, $[\mathbf{a}'',\mathbf{a}]$ and $[\mathbf{a}'',\mathbf{a}']$.

THEOREM 12·1. *Let Φ and Ψ be subsets of a shape range W and F be an absolutely free group of rank at least* 3. *Then $\hat{W}_\Phi(F) \geqslant \hat{W}_\Psi(F)$ if and only if $\Phi \preccurlyeq \Psi$.*

Proof. By virtue of theorem 3·1 (iv) it is sufficient to prove that if $\Phi \not\preccurlyeq \Psi$ then $\hat{W}_\Phi(F) \ngeqslant \hat{W}_\Psi(F)$. Suppose then that $\Phi \not\preccurlyeq \Psi$: then there exists $\psi \in \Psi$ such that

$$\phi \in \Phi \Rightarrow \phi \not\preccurlyeq \psi,$$

that is, $\Phi \not\preccurlyeq \{\psi\}$. Let $\rho: \mathbf{A} \to F$ be a free description of F. Then the number of generators of $\mathbf{A}$ is at least 3 so, by lemma 12·3, there exists a basic commutator $\mathbf{c} \in \mathbf{A}$ such that $\hat{\sigma}(\mathbf{c}) = \sigma(\mathbf{c}) = \psi$. Then $\Sigma(\mathbf{c}) = \{\psi\} \not\succcurlyeq \Phi$ and thus $\mathbf{c}\rho \neq \mathbf{1}\rho$ modulo $\hat{W}_\Phi(F)$ by the Basis Theorem (theorem 9·1 (*c*) (ii)), that is, $\mathbf{c}\rho \notin \hat{W}_\Phi(F)$. But $\Sigma(\mathbf{c}) = \{\psi\} \succcurlyeq \Psi$ so $\mathbf{c}\rho \in \hat{W}_\Psi(F)$. Then

$$\mathbf{c}\rho \in \hat{W}_\Psi(F) - \hat{W}_\Phi(F)$$

so $\hat{W}_\Phi(F) \ngeqslant \hat{W}_\Psi(F)$.

This theorem establishes the facts concerning distinctness of shape subgroups mentioned at the beginning of this section.

As regards whether rank 3 is best possible, clearly one generator is not enough to do more than distinguish the whole group ($W_1(F)$, $\hat{W}_1(F)$ and $\hat{W}_\Phi(F)$ whenever $1 \in \Phi$) from the trivial subgroup (all other subgroups). For some particular shape ranges, such as N^-, two generators are enough. An example for which two generators are not enough must use the language and results of chapter III. If $K = (k_i)_{i=1}^\infty$ is the sequence $k_i = 2$ (all i), then

$$\hat{Q}_{\delta_2}(F_2) = Q_{\delta_2}(F_2) = \delta^2(F_2) \quad \text{and} \quad \hat{Q}_{\delta_2+1}(F_2) = Q_{\delta_2+1}(F_2) = \delta^2(F_2) \cap \gamma_5(F_2).$$

But it is well known that for a two-generator group these subgroups are the same.

13. *Partial collection*

The 'non-domination' restriction appearing in lemma 9·4 and consequently in theorem 9·1 (A) (iii), (B) (iii) and (C) (iii) is disquieting: it means that there is no guarantee that an arbitrary expression can be collected at all. That this restriction is real and not just due to an inadequacy in the method of proof is demonstrated in appendix I where it is shown that if β finely dominates α and provided that the rank of the absolutely free group F is at least 3, there exists an element in $W_\alpha(F)$ which cannot be described by a basic expression modulo $W_\beta(F)$ at all.

In default of this, a property which would be a good second-best is as follows.

DEFINITION 13·1. *A shape range W is* partially collectable *if, for any $\Phi \subseteq W$, any description $\rho: \mathbf{A} \to G$ of a group G and any $x \in G - \hat{W}_\Phi(G)$, there exists a non-negative integer c and a basic expression $\mathbf{x} \in \hat{\mathbf{B}}_\Phi \cap \mathbf{B}_{(c)}$ such that $\mathbf{x} \neq \mathbf{1}$ and $\mathbf{x}\rho = x$* modulo *$\hat{W}_\Phi(G) \cdot \gamma_{c+1}(G)$.*

In chapter III it will be shown that polyweight ranges have this property and its importance will emerge as it is used in proofs. Meanwhile, some elementary properties are exhibited for later use.

LEMMA 13·1. *Let W be a shape range. Then the following three properties are equivalent.*

(i) *W is partially collectable.*

(ii) *For any* $\Phi \subseteq W$ *the relatively free group* $F(\hat{\mathfrak{W}}_\Phi)$ *of any rank is residually nilpotent.*

(iii) *For any* $\Phi \subseteq W$ *and any prime p the relatively free group* $F(\hat{\mathfrak{W}}_\Phi)$ *of any rank is residually a finite p-group.*

Proof. The equivalence of (i) and (ii) follows immediately from the Basis Theorem. Writing $G = F(\hat{\mathfrak{W}}_\Phi)$, the same theorem implies that $G/\gamma_c(G)$ is torsion-free, which makes the equivalence of (ii) and (iii) obvious.

LEMMA 13·2. *Let W be a partially collectable shape range. Then*

(i) *For any* $\alpha \in W$, *any description* $\rho: \mathbf{A} \to G$ *of a group G and any* $x \in G - W_\alpha(G)$, *there exists a non-negative integer c and a basic expression* $\mathbf{x} \in \mathbf{B}_\alpha \cap \mathbf{B}_{(c)}$ *such that*

$$\mathbf{x} \neq 1 \quad \textit{and} \quad \mathbf{x}\rho = x \text{ modulo } W_\alpha(G) \, . \, \gamma_{c+1}(G).$$

(ii) *For any* $\alpha \in W$, *any description* $\rho: \mathbf{A} \to G$ *of a group G and any* $x \in G - \hat{W}_\alpha(G)$, *there exists a non-negative integer c and a basic expression* $\mathbf{x} \in \hat{\mathbf{B}}_\alpha \cap \mathbf{B}_{(c)}$ *such that*

$$\mathbf{x} \neq 1 \quad \textit{and} \quad \mathbf{x}\rho = x \text{ modulo } \hat{W}_\alpha(G) \, . \, \gamma_{c+1}(G).$$

Proof. It has been observed that fine and coarse shape subgroups may be expressed as shape set subgroups. The result follows immediately.

The next lemma is a straightforward illustration of the use of this property.

LEMMA 13·3. *Let W be a partially collectable shape range,* Φ *and* Ψ *subsets of W and F an absolutely free group. Then*

$$\hat{W}_\Phi(F) \cap \hat{W}_\Psi(F) = \hat{W}_{\Phi \vee \Psi}(F).$$

Proof. By virtue of theorem 3·1 (vi) it is sufficient to prove that if $x \notin \hat{W}_{\Phi\vee\Psi}(F)$ then $x \notin \hat{W}_\Phi(F) \cap \hat{W}_\Psi(F)$. Suppose then that $x \notin W_{\Phi\vee\Psi}(F)$. Then there exists a non-negative integer c and a basic expression $\mathbf{x} \in \hat{\mathbf{B}}_{\Phi\vee\Psi} \cap \mathbf{B}_{(c)}$ such that

$$\mathbf{x} \neq 1 \quad \text{and} \quad \mathbf{x}\rho = x \text{ modulo } \hat{W}_{\Phi\vee\Psi}(F) \, . \, \gamma_c(F).$$

Since $\mathbf{x} \neq 1$ it may be written in the form

$$\mathbf{x} = \mathbf{b}_1^{\alpha_1} \mathbf{b}_2^{\alpha_2} \ldots \mathbf{b}_k^{\alpha_k} \quad (k \geqslant 1)$$

and, since $\mathbf{x} \in \hat{\mathbf{B}}_{\Phi\vee\Psi}$, $\Sigma(\mathbf{b}_1) \not\geqslant \Phi \vee \Psi$ so that either $\Sigma(\mathbf{b}_1) \not\geqslant \Phi$ or $\Sigma(\mathbf{b}_1) \not\geqslant \Psi$. It may be assumed without loss of generality that $\Sigma(\mathbf{b}_1) \not\geqslant \Phi$. Then by the Basis Theorem (theorem 9·1 (D) (iii)), $\mathbf{x}\rho \notin \hat{W}_\Phi(F) \, . \, \gamma_c(F)$ and thus $x \notin \hat{W}_\Phi(F) \, . \, \gamma_c(F)$ since $W_{\Phi\vee\Psi}(F) \leqslant W_\Phi(F)$. Hence $x \notin \hat{W}_\Phi(F) \geqslant \hat{W}_\Phi(F) \cap \hat{W}_\Psi(F)$.

CHAPTER III. POLYWEIGHTS

14. *The shape*

DEFINITION 14·1. *Let* $K = (k_i)_{i=1}^\infty$ *be a sequence of integers, each* $\geqslant 2$. *For each non-negative integer r, let* K_r *be the finite sequence* $K_r = (k_i)_{i=1}^r$; *in particular,* K_0 *is the 'empty' sequence. For any group G and each* K_r *the subgroup* $P_{K_r}(G)$ *is defined recursively over r by*

$$P_{K_0}(G) = G, \qquad P_{K_r}(G) = \gamma_{k_r}(P_{K_{r-1}}(G)) \quad (r > 0).$$

The resulting series $P_{K_0}(G), P_{K_1}(G), \ldots$ *will be called the* polycentral series *of G of* type *K*.

A group G is polynilpotent of type K_r *if* $P_{K_r}(G) = \{1\}$. *The class of all such groups is a variety, the* polynilpotent variety *of* type K_r, *denoted* $\mathfrak{P}_{K_r}$.

For each sequence K a shape range Q^K will now be defined which will have the property that all the varieties $\mathfrak{P}_{K_r}$ will be among the varieties $\mathfrak{Q}^K_\alpha$.

Definition 14·2. *Let $K = (k_i)_{i=1}^{\infty}$ be a sequence of integers, each $\geqslant 2$. Then*

(A) *Q^K is the set of all functions $\phi: \omega \to \omega$ satisfying*

(i) $\phi(j-1) \geqslant k_j \phi(j)$ (*for all $j \geqslant 1$*),

(ii) $\phi(j-1) \geqslant k_j \Rightarrow \phi(j) \geqslant 1$ (*for all $j \geqslant 1$*) *and*

(iii) $\phi(0) \geqslant 1$,

together with an extra element ∞. The function $1 \in Q^K$ is defined by $1(0) = 1$, $1(j) = 0$ (for all $j \geqslant 1$).

It follows from (i) *and* (ii), *since each $k_j \geqslant 2$, that for each function $\phi \neq \infty$ there exists an integer d_ϕ, the* depth *of ϕ, such that $\phi(j) \neq 0 \Leftrightarrow j \leqslant d_\phi$. The depth d_∞ of ∞ is defined to be ∞.*

(B) *Addition is defined on Q^K as follows: if ϕ and ψ are functions (that is, $\neq \infty$) then*

(i) *if $d_\phi = d_\psi$ ($= d$ say) and $\phi(d) + \psi(d) \geqslant k_{d+1}$ then*

$$(\phi+\psi)(j) = \phi(j)+\psi(j) \quad (j \neq d+1),$$
$$= 1 \quad (j = d+1),$$

(ii) *otherwise*
$$(\phi+\psi)(j) = \phi(j)+\psi(j) \quad (\text{for all } j).$$

Addition is extended to encompass ∞ by

$$\phi+\infty = \infty+\phi = \infty \quad (\text{for all } \phi \in Q^K).$$

(C) *The* fine order $\leqslant$ *on Q^K is defined lexicographically from the right: if ϕ and ψ are functions ($\neq \infty$) then $\phi < \psi$ if and only if there exists $j_0 \in \omega$ such that*

$$\phi(j_0) < \psi(j_0) \quad \text{and} \quad j > j_0 \Rightarrow \phi(j) = \psi(j).$$

This order is extended to encompass ∞ by: if ϕ is a function then $\phi < \infty$.

(D) *The* coarse order $\preccurlyeq$ *on Q^K is defined: if ϕ and ψ are functions ($\neq \infty$) then $\phi \preccurlyeq \psi$ if and only if $\phi(j) \leqslant \psi(j)$ for all $j \in \omega$. For any function ϕ, $\phi \prec \infty$.*

The set Q^K together with this addition and orders is the polyweight range *of* type K. *Subject to the proof given in this section that Q^K is in fact a shape range, the associated fine shape, coarse shape and shape set are called the* fine polyweight, coarse polyweight *and* polyweight set *of type K and denoted π^K, $\hat{\pi}^K$ and Π^K respectively.*

The remainder of this section will be occupied with a proof that Q^K is indeed a shape range. For clarity this proof will be broken down into several lemmas.

It follows immediately from part (A) of the definition that Q^K is closed under addition.

For use in the following lemmas, some simple properties of the orders should be remarked:

$$\phi \preccurlyeq \psi \Rightarrow \phi \leqslant \psi, \quad d_\phi < d_\psi \Rightarrow \phi < \psi, \quad \phi \leqslant \psi \Rightarrow d_\phi \leqslant d_\psi,$$
$$(\phi \wedge \psi)(j) = \min\{\phi(j), \psi(j)\} \quad \text{and} \quad (\phi \vee \psi)(j) = \max\{\phi(j), \psi(j)\}.$$

Lemma 14·1. *Q^K is (fully) well-ordered by the fine order $\leqslant$. The least element is the function 1, the greatest, the element ∞.*

Proof. A non-empty subset X of Q^K either consists of the element ∞ alone, in which case it has a least element trivially, or else it contains a function $\phi \neq \infty$. But then the set of all functions $\leqslant \phi$ in X consists of functions whose support is a subset of the (finite) support of ϕ. Since the order is lexicographic the lemma follows.

LEMMA 14·2. *If* α_1, α_2 *and* β *are functions* $(\neq \infty)$ *in* Q^K *and* $\alpha_1 < \alpha_2$, *then* $\alpha_1+\beta < \alpha_2+\beta$.

Proof. Since $\alpha_1 < \alpha_2$, $d_{\alpha_1} \leqslant d_{\alpha_2}$ and there exists $j_0 \in \omega$ such that $\alpha_1(j_0) < \alpha_2(j_0)$ and $j > j_0 \Rightarrow \alpha_1(j) = \alpha_2(j)$. Four cases now arise: case 1, $d_{\alpha_1} = d_{\alpha_2} = d_\beta$, case 2, $d_\beta \neq d_{\alpha_1} = d_{\alpha_2}$, case 3, $d_\beta = d_{\alpha_1} < d_{\alpha_2}$ and case 4, $d_\beta \neq d_{\alpha_1} < d_{\alpha_2}$.

Case 1, $d_{\alpha_1} = d_{\alpha_2} = d_\beta$ $(= d$ say). Then $j_0 \leqslant d$. Now

$$(\alpha_1+\beta)(d+1) = 0 \text{ or } 1 \quad \text{and} \quad (\alpha_2+\beta)(d+1) = 0 \text{ or } 1$$

and if $(\alpha_1+\beta)(d+1) = 1$ then $\alpha_1(d)+\beta(d) \geqslant k_{d+1}$. But $j_0 \leqslant d$ so $\alpha_1(d) \leqslant \alpha_2(d)$. Thus $\alpha_2(d)+\beta(d) \geqslant k_{d+1}$ and then $(\alpha_2+\beta)(d+1) = 1$ also and so in any case,

$$(\alpha_1+\beta)(d+1) \leqslant (\alpha_2+\beta)(d+1).$$

Thus $j > j_0 \Rightarrow (\alpha_1+\beta)(j) \leqslant (\alpha_2+\beta)(j)$ and, since $j_0 \leqslant d$,

$$(\alpha_1+\beta)(j_0) = \alpha_1(j_0)+\beta(j_0) < \alpha_2(j_0)+\beta(j_0) = (\alpha_2+\beta)(j_0).$$

This means that $\alpha_1+\beta < \alpha_2+\beta$.

Case 2, $d_\beta \neq d_{\alpha_1} = d_{\alpha_2}$. Then

$$j > j_0 \Rightarrow (\alpha_1+\beta)(j) = \alpha_1(j)+\beta(j) = \alpha_2(j)+\beta(j) = (\alpha_2+\beta)(j)$$

and
$$(\alpha_1+\beta)(j_0) = \alpha_1(j_0)+\beta(j_0) < \alpha_2(j_0)+\beta(j_0) = (\alpha_2+\beta)(j_0).$$

Thus $\alpha_1+\beta < \alpha_2+\beta$.

Case 3, $d_\beta = d_{\alpha_2} < d_{\alpha_2}$. Let $d = d_\beta = d_{\alpha_1}$. Then $j_0 = d_{\alpha_2}$. For $j > d+1$, $(\alpha_1+\beta)(j) = 0$ and $(\alpha_2+\beta)(j) = \alpha_2(j) \geqslant 0$; also $(\alpha_1+\beta)(d+1) = 0$ or 1 and $(\alpha_2+\beta)(d+1) = \alpha_2(d+1) \geqslant 1$. Thus $j \geqslant d+1 \Rightarrow (\alpha_1+\beta)(j) \leqslant (\alpha_2+\beta)(j)$. But $\alpha_1(d) < k_{d+1}$ and $\alpha_2(d) \geqslant k_{d+1}$ so

$$(\alpha_1+\beta)(d) < (\alpha_2+\beta)(d).$$

Thus $\alpha_1+\beta < \alpha_2+\beta$.

Case 4, $d_\beta \neq d_{\alpha_1} < d_{\alpha_2}$. Then

$$j > j_0 \Rightarrow (\alpha_1+\beta)(j) = \alpha_1(j)+\beta_1(j) = \alpha_2(j)+\beta(j) \leqslant (\alpha_2+\beta)(j)$$

and
$$(\alpha_1+\beta)(j_0) = \alpha_1(j_0)+\beta(j_0) < \alpha_2(j_0)+\beta(j_0) = (\alpha_2+\beta)(j_0)$$

and again $\alpha_1+\beta < \alpha_2+\beta$.

LEMMA 14·3. *If* α_1, α_2 *and* β *are functions* $(\neq \infty)$ *in* Q^K *and* $\alpha_1 \prec \alpha_2$, *then* $\alpha_1+\beta \prec \alpha_2+\beta$.

Proof. By the definition of the coarse order, for all $j \in \omega$, $\alpha_1(j) \leqslant \alpha_2(j)$ and there exists $j_0 \in \omega$ such that $\alpha_1(j_0) < \alpha_2(j_0)$.

First, suppose that $d_{\alpha_1} = d_\beta$ and write $d = d_{\alpha_1} = d_\beta$. Then $(\alpha_1+\beta)(d+1) = 0$ or 1 and if $(\alpha_1+\beta)(d+1) = 1$ then $\alpha_1(d)+\beta(d) \geqslant k_{d+1}$. But then $\alpha_2(d)+\beta(d) \geqslant k_{d+1}$ so $(\alpha_2+\beta)(d+1) \geqslant 1$ Thus in any case $(\alpha_1+\beta)(d+1) \leqslant (\alpha_2+\beta)(d+1)$. But, for $j \neq d+1$,

$$(\alpha_1+\beta)(j) = \alpha_1(j)+\beta(j) \leqslant \alpha_2(j)+\beta(j) \leqslant (\alpha_2+\beta)(j),$$

so $(\alpha_1+\beta)(j) \leqslant (\alpha_2+\beta)(j)$ for all $j \in \omega$. If $j_0 \neq d+1$ then

$$(\alpha_1+\beta)(j_0) = \alpha_1(j_0)+\beta(j_0) < \alpha_2(j_0)+\beta(j_0) \leqslant (\alpha_2+\beta)(j_0)$$

and if $j_0 = d+1$ then $\alpha_2(d+1) \geqslant 1$ so $\alpha_2(d) \geqslant k_{d+1} > \alpha_1(d)$ and $(\alpha_1+\beta)(d) < (\alpha_2+\beta)(d)$. Thus in any case $\alpha_1+\beta \prec \alpha_2+\beta$.

Now suppose that $d_{\alpha_1} \neq d_\beta$. Then, for all $j \in \omega$,

$$(\alpha_1+\beta)(j) = \alpha_1(j)+\beta(j) \leqslant \alpha_2(j)+\beta(j) \leqslant (\alpha_2+\beta)(j)$$

and
$$(\alpha_1+\beta)(j_0) = \alpha_1(j_0)+\beta(j_0) < \alpha_2(j_0)+\beta(j_0) \leqslant (\alpha_2+\beta)(j_0).$$

Lemma 14·4. *Suppose α, β and γ are functions ($\neq \infty$) in Q^K and $\alpha \leqslant \beta \leqslant \gamma$. Then*

(i) *$\gamma+\beta+\alpha = \gamma+\alpha+\beta \leqslant \beta+\alpha+\gamma$ and*

(ii) *$\gamma+\beta+\alpha = \gamma+\alpha+\beta \prec \beta+\alpha+\gamma$ if and only if $d_\alpha = d_\beta < d_{\alpha+\beta} \leqslant d_\gamma$.*

Proof. Four cases are considered separately: case 1, $d_\alpha = d_\beta = d_\gamma$, case 2, $d_\alpha = d_\beta < d_\gamma$, case 3, $d_\alpha < d_\beta = d_\gamma$ and case 4, $d_\alpha < d_\beta < d_\gamma$.

Case 1, $d_\alpha = d_\beta = d_\gamma$ ($= d$ say). Suppose that $\alpha(d)+\beta(d)+\gamma(d) < k_{d+1}$. Then, for all j,

$$(\gamma+\beta+\alpha)(j) = (\gamma+\alpha+\beta)(j) = (\beta+\alpha+\gamma)(j) = \alpha(j)+\beta(j)+\gamma(j).$$

Otherwise $\alpha(d)+\beta(d)+\gamma(d) \geqslant k_{d+1}$ and then, for $j \neq d+1$,

$$(\gamma+\beta+\alpha)(j) = (\gamma+\alpha+\beta)(j) = (\beta+\alpha+\gamma)(j) = \alpha(j)+\beta(j)+\gamma(j)$$

and
$$(\gamma+\beta+\alpha)(d+1) = (\gamma+\alpha+\beta)(d+1) = (\beta+\alpha+\gamma)(d+1) = 1.$$

Case 2, $d_\alpha = d_\beta < d_\gamma$. Let $d = d_\alpha = d_\beta$. Then $d_\beta \neq d_\gamma$ so $(\gamma+\beta)(j) = \beta(j)+\gamma(j)$ for all j. Then $d_{\gamma+\beta} = d_\gamma \neq d_\alpha$ so

$$(\gamma+\beta+\alpha)(j) = \alpha(j)+\beta(j)+\gamma(j) \quad \text{for all } j.$$

Similarly,
$$(\gamma+\alpha+\beta)(j) = \alpha(j)+\beta(j)+\gamma(j) \quad \text{for all } j.$$

Now $d_\alpha = d_\beta = d$ so $d_{\beta+\alpha} = d$ or $d+1$. If $d_{\beta+\alpha} = d$ then $d_{\beta+\alpha} \neq d_\gamma$ so

$$(\beta+\alpha+\gamma)(j) = \alpha(j)+\beta(j)+\gamma(j) \quad \text{for all } j.$$

If $d_{\beta+\alpha} = d+1$ then $d_\alpha = d_\beta < d_{\alpha+\beta} \leqslant d_\gamma$ and

$$\begin{aligned}(\beta+\alpha)(j) &= \alpha(j)+\beta(j) \quad && (j \neq d+1),\\ &= 1 && (j = d+1),\end{aligned}$$

and then
$$\begin{aligned}(\beta+\alpha+\gamma)(j) &= \alpha(j)+\beta(j)+\gamma(j) \quad && (j \neq d+1),\\ &= \gamma(j)+1 && (j = d+1).\end{aligned}$$

Then, since $\alpha(d+1) = \beta(d+1) = 0$, $\beta+\alpha+\gamma \succ \gamma+\beta+\alpha = \gamma+\alpha+\beta$.

Case 3, $d_\alpha < d_\beta = d_\gamma$. Let $d = d_\beta = d_\gamma$. Then there are two subcases:

case 3·1, $\beta(d)+\gamma(d) \geqslant k_{d+1}$ and case 3·2, $\beta(d)+\gamma(d) < k_{d+1}$.

Case 3·1, $\beta(d)+\gamma(d) \geqslant k_{d+1}$. Then

$$\begin{aligned}(\gamma+\beta)(j) &= \beta(j)+\gamma(j) \quad && (j \neq d+1)\\ &= 1 && (j = d+1)\end{aligned}$$

and so $d_{\gamma+\beta} = d+1 \neq \alpha$. Thus

$$\begin{aligned}(\gamma+\beta+\alpha)(j) &= \alpha(j)+\beta(j)+\gamma(j) \qquad (j \neq d+1)\\ &= 1 \qquad\qquad\qquad\qquad\quad (j = d+1)\end{aligned}$$

since $\alpha(d+1) = 0$. Now $d_\alpha \neq d_\gamma$ so $(\gamma+\alpha)(j) = \alpha(j)+\gamma(j)$ for all j. Thus $d_{\gamma+\alpha} = d_\gamma = d$. But then $(\gamma+\alpha)(d)+\beta(d) = \gamma(d)+\beta(d) \geqslant k_{d+1}$ so

$$\begin{aligned}(\gamma+\alpha+\beta)(j) &= \alpha(j)+\beta(j)+\gamma(j) \qquad (j \neq d+1)\\ &= 1 \qquad\qquad\qquad\qquad\quad (j = d+1)\end{aligned}$$

so that $\gamma+\alpha+\beta = \gamma+\beta+\alpha$. Similarly, $\beta+\alpha+\gamma = \gamma+\beta+\alpha$.

Case 3·2, $\beta(d)+\gamma(d) < k_{d+1}$. Then $(\gamma+\beta)(j) = \beta(j)+\gamma(j)$ for all j. Then $d_{\gamma+\beta} = d \neq d_\alpha$ so that

$$(\gamma+\beta+\alpha)(j) = \alpha(j)+\beta(j)+\gamma(j) \quad \text{for all } j.$$

Now $d_\gamma = d \neq d_\alpha$ so that $\quad (\gamma+\alpha)(j) = \alpha(j)+\gamma(j) \quad$ for all j.

Thus $d_{\gamma+\alpha} = d = d_\beta$. But $(\gamma+\alpha)(d)+\beta(d) = \gamma(d)+\beta(d) < k_{d+1}$ so that

$$(\gamma+\alpha+\beta)(j) = \alpha(j)+\beta(j)+\gamma(j) \quad \text{for all } j.$$

Similarly, $\quad (\beta+\alpha+\gamma)(j) = \alpha(j)+\beta(j)+\gamma(j) \quad$ for all j.

Case 4, $d_\alpha < d_\beta < d_\gamma$. Then

$$(\gamma+\beta+\alpha)(j) = (\gamma+\alpha+\beta)(j) = (\beta+\alpha+\gamma)(j) = \alpha(j)+\beta(j)+\gamma(j) \quad \text{for all } j.$$

The condition $\alpha \leqslant \beta \leqslant \gamma$ may be removed from part (ii) of the lemma as follows.

Corollary. *Suppose* α, β *and* γ *are functions* ($\neq \infty$) *in* Q^K. *Then*

(i) $\gamma+\beta+\alpha$ *and* $\beta+\alpha+\gamma$ *are comparable* (*under the coarse order*) *and*

(ii) $\gamma+\beta+\alpha \prec \beta+\alpha+\gamma$ *if and only if* $d_\alpha = d_\beta < d_{\alpha+\beta} = d_\gamma$.

Proof. The various order relationships that may exist between α, β and γ are checked separately. If $\alpha \leqslant \beta \leqslant \gamma$ the result is given by the lemma. Suppose $\alpha \leqslant \gamma \leqslant \beta$. Then

$$\begin{aligned}\gamma+\beta+\alpha &= \beta+\gamma+\alpha \quad \text{(by commutativity of +)},\\ &= \beta+\alpha+\gamma \quad \text{(by the lemma)}.\end{aligned}$$

But $\beta \geqslant \gamma$ so $d_\beta \nless d_\gamma$. Now suppose $\beta \leqslant \alpha \leqslant \gamma$. Then

$$\gamma+\beta+\alpha \prec \alpha+\beta+\gamma$$

if and only if $d_\alpha = d_\beta < d_{\alpha+\beta} \leqslant d_\gamma$, by the lemma, and

$$\alpha+\beta+\gamma = \beta+\alpha+\gamma \quad \text{(by commutativity of +)}.$$

Proofs for the other three cases are similar.

Definition 14·3. (i) *With the notation of definition* 14·2, *for each non-negative integer* r, *a function* $\delta_r^K \in Q^K$ *is defined by*

$$\begin{aligned}\delta_r^K(j) &= 0 \qquad\qquad\qquad (j > r)\\ &= 1 \qquad\qquad\qquad (j = r)\\ &= k_{j+1}\,\delta_r^K(j+1) \quad (j < r)\end{aligned}$$

the definition of $\delta_r^K(j)$ *for* $j \leqslant r$ *being recursive over* $r-j$.

(ii) *For any positive integer n, a function* $n\delta_r^K \in Q^K$ *is defined recursively*

$$1\delta_r^K = \delta_r^K,$$
$$n\delta_r^K = (n-1)\,\delta_r^K + \delta_r^K \quad (n > 1).$$

That δ_r^K is indeed a member of Q^K follows from this definition and definition 14·2 (A). That $n\delta_r^K \in Q^K$ then follows from the closure of Q^K under addition.

LEMMA 14·5

(i) *For* $0 \leqslant j < r$,

$$\delta_r^K(j) = \prod_{i=j+1}^{r} k_i.$$

(ii) *For any positive integer n,* $n\delta_r^K$ *may be defined alternatively by*

$$\begin{aligned} n\delta_r^K(j) &= 0 && (j > r+1)\\ &= \begin{cases} 0 & \text{if } n < k_{r+1} \\ 1 & \text{if } n \geqslant k_{r+1} \end{cases} && (j = r+1)\\ &= n && (j = r)\\ &= k_{j+1}\, n\delta_r^K(j+1) && (j < r). \end{aligned}$$

(iii) *For any function* $\phi \in Q^K$, $\quad \phi = \bigvee_{r \leqslant d_\phi} \phi(r)\,\delta_r^K.$

(iv) *For* $0 \leqslant i \leqslant j \leqslant r$, $\quad \delta_r^K(i) = \delta_r^K(j)\,.\,\delta_j^K(i).$

Proof. These all follow immediately from the definition, part (iii) by using definition 14·2 (A) (i).

LEMMA 14·6. $Q^K - \{\infty\}$ *is generated by* 1 *under addition.*

Proof. By virtue of lemmas 14·1 and 14·2, it is sufficient to show that for each $\phi \in Q^K$ other than 1 or ∞, there exist $\psi_1, \psi_2 \in Q^K$ such that $\phi = \psi_1 + \psi_2$. Three cases are considered, depending on the values of $\phi(d)$ and $\phi(d-1)$ where $d = d_\phi$.

Case 1, $\phi(d) \geqslant d$. Let $\psi_1 = \delta_d^K$ and define ψ_2 by $\psi_2(j) = \phi(j) - \psi_1(j)$ for all j. Then for all $j \geqslant 1$,

$$\psi_2(j-1) = \phi(j-1) - \psi_1(j-1) \geqslant k_j \phi(j) - k_j \psi(j) = k_j \psi_2(j)$$

and $\psi_2(j-1) \geqslant k_j \Rightarrow j \leqslant d$ (since $\phi(d) < k_{d+1}) \Rightarrow \psi_2(j) \geqslant 1$. Trivially $\psi_2(0) \geqslant 1$. Hence $\psi_2 \in Q^K$. But now $d_{\psi_1} = d_{\psi_2} = d$ and $\psi_1(d) + \psi_2(d) = \phi(d) < k_{d+1}$ so

$$(\psi_1 + \psi_2)\,(j) = \psi_1(j) + \psi_2(j) = \phi(j)$$

for all j.

Case 2, $\phi(d) = 1$ and $\phi(d-1) > k_d$. Write $\psi_1 = \delta_{d-1}^K$ and define ψ_2 by $\psi_2(j) = \phi(j) - \psi_1(j)$ for all j; the argument now goes as before.

Case 3, $\phi(d) = 1$ and $\phi(d-1) = k_d$. Write $\psi_1 = \delta_{d-1}^K$ and define ψ_2 by $\psi_2(j) = \phi(j) - \psi_1(j)$ for all $j \neq d$, $\psi_2(d) = 0$; the argument goes as before.

THEOREM 14·1. *With the notation of definition* 14·2, Q^K *is a shape range.*

Proof. The parts of definition 2·1 are checked.

(i) The fine order $\leqslant$ (fully) well-orders Q^K, 1 is the least element and ∞ the greatest by lemma 14·1. That the coarse order $\preccurlyeq$ is a lattice order follows immediately from its definition.

(ii) Q^K is closed under addition and $Q^K-\{\infty\}$ is generated by 1 under addition by lemma 14·6. That addition is commutative follows immediately from its definition.

(iii) By the definition of addition, $\alpha, \beta \neq \infty \Rightarrow \alpha+\beta \neq \infty$ and $\alpha+\infty = \infty$. Now suppose $\alpha_1 < \alpha_2$ and $\beta \neq \infty$. If $\alpha_2 = \infty$ then $\alpha_1+\beta < \infty = \alpha_2+\beta$ and if $\alpha_2 \neq \infty$ then $\alpha_1+\beta < \alpha_2+\beta$ by lemma 14·2.

(iv) If $\alpha \neq \infty$, then, by the definition of addition, $\alpha \prec \alpha+\beta$. That

$$\alpha_1 \prec \alpha_2 \quad \text{and} \quad \beta \neq \infty \Rightarrow \alpha_1+\beta \prec \alpha_2+\beta$$

follows as above from lemma 14·3.

(v) Suppose $\alpha \leqslant \beta \leqslant \gamma$. If any one is ∞ then $\gamma+\beta+\alpha = \gamma+\alpha+\beta = \beta+\alpha+\gamma = \infty$, otherwise $\gamma+\beta+\alpha = \gamma+\alpha+\beta \leqslant \beta+\alpha+\gamma$ by lemma 14·4 (i).

Lemma 14·7. *Let* $\mathbf{A}$ *be an algebra of expressions,* $\mathbf{x} \in \mathbf{A}$ *and* $\hat{\pi}: \mathbf{A} \to Q^K$, $\pi: \mathbf{A} \to Q^K$ *the coarse and fine polyweights associated with* Q^K. *Then, writing* $\phi = \pi(\mathbf{x})$ *and* $\hat{\phi} = \hat{\pi}(\mathbf{x})$,

$$\phi(0) = \hat{\phi}(0) = \mathrm{wt}\,(\mathbf{x}).$$

Proof is by an easy induction over the height of $\mathbf{x}$.

15. *Partial collectability of polyweights*

Suppose W and W' are two arbitrary shape ranges. Then for any $\alpha \in W$ the set $\hat{\mathbf{W}}_\alpha$ (definition 3·1) is a subalgebra of $\mathbf{A}$ and hence a describable algebra. Consequently, for any $\Phi \subseteq W'$, the set $\hat{W}'_\Phi(\hat{\mathbf{W}}_\alpha)$ is defined (definition 3·2). Further, this is a fully invariant subalgebra of $\mathbf{A}$ and defines a product variety of groups, since for any description $\rho: \mathbf{A} \to G$ of a group G, $(\hat{W}'_\Phi(\hat{\mathbf{W}}_\alpha))\rho = \hat{W}'_\Phi(\hat{W}_\alpha(G))$.

Lemma 15·1. *Suppose* $K = (k_i)_{i=1}^{\infty}$ *is any sequence of integers, each* $\geqslant 2$ *and* $K' = (k'_i)_{i=1}^{\infty}$ *is defined by* $k'_i = k_{i+1}$ *(for all* $i \geqslant 1$*). Write* $Q = Q^K$ *and* $Q' = Q^{K'}$. *Then for any function* $\phi \in Q$ *with the property* $\phi(0) = k_1\phi(1)$ *and any group* G, $\hat{Q}_\phi(G) = \hat{Q}'_{\phi'}(\gamma_{k_1}(G))$, *where* ϕ' *is the function defined by* $\phi'(j) = \phi(j+1)$ *for all* j.

Proof. First it is necessary to observe that $\phi' \in Q'$ so that $\hat{Q}'_{\phi'}(\gamma_{k_1}(G))$ has meaning. Let $\rho: \mathbf{A} \to G$ be any description and write $\hat{\pi} = \hat{\pi}^K$ and $\hat{\pi}' = \hat{\pi}^{K'}$. For any $\alpha \in Q$ such that $d_\alpha \geqslant 1$, define α' by $\alpha'(j) = \alpha(j+1)$ for all $j \in \omega$: then clearly $\alpha' \in Q'$ and for any $\alpha, \beta \in Q$ of depth $\geqslant 1$, $(\alpha+\beta)' = \alpha'+\beta'$.

First is is shown that $\hat{Q}_\phi(G) \leqslant \hat{Q}'_{\phi'}(\gamma_{k_1}(G))$. This is accomplished by proving that, for $\mathbf{x} \in \mathbf{A}$,

$$\hat{\pi}(\mathbf{x}) \geqslant \alpha \quad \text{and} \quad d_\alpha \geqslant 1 \Rightarrow \mathbf{x}\rho \in \hat{Q}'_{\phi'}(\gamma_{k_1}(G)),$$

by induction over the height of $\mathbf{x}$. If $\mathrm{ht}\,(\mathbf{x}) = 1$ then either $\mathbf{x} = \mathbf{1}$ in which case

$$\mathbf{x}\rho = 1 \in Q'_{\phi'}(\gamma_{k_1}(G))$$

trivially, or else $\mathbf{x} = \mathbf{g}_i \in \mathbf{G}$ in which case $d_{\hat{\pi}(x)} = 0$ and the implication is satisfied vacuously. Now suppose that $\mathrm{ht}\,(\mathbf{x}) > 1$ and the implication is true for all expressions of smaller height. Then there are three possibilities: if $\mathbf{x} = \mathbf{u}^{-1}$ or $\mathbf{x}_1\mathbf{x}_2$ the result follows immediately from the fact that $\hat{Q}'_{\phi'}(\gamma_{k_1}(G))$ is a subgroup and if $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2]$, write $\hat{\pi}(\mathbf{x}) = \psi$ (so that $\psi \geqslant \alpha$), $\hat{\pi}(\mathbf{x}_1) = \psi_1$ and $\hat{\pi}(\mathbf{x}_2) = \psi_2$. Then $\psi = \psi_1+\psi_2$ and four cases arise.

Case 1, $d_{\psi_1} = d_{\psi_2} = 0$. Then $\psi(1) = 1$ since $\psi \geqslant \alpha$ and $d_\alpha \geqslant 1$ and so $\psi' = 1$ and $\psi(0) \geqslant k_1$. Thus $\text{wt}(\mathbf{x}) = \psi(0) \geqslant k_1$ by lemma 14·7 and so

$$\mathbf{x}\rho \in \gamma_{k_1}(G) = \hat{Q}'_1(\gamma_{k_1}(G)) = \hat{Q}'_{\alpha'}(\gamma_{k_1}(G))$$

for $\alpha \leqslant \psi$ and $\psi' = 1$ so that $\alpha' = 1$.

Case 2, $d_{\psi_1} = 0$, $d_{\psi_2} \geqslant 1$. Then $\psi'_2 = \psi' \geqslant \alpha'$ and so, by the inductive hypothesis, $\mathbf{x}_2\rho \in \hat{Q}'_{\alpha'}(\gamma_{k_1}(G))$; but this is a normal subgroup of G and hence also contains $[\mathbf{x}_1\rho, \mathbf{x}_2\rho] = \mathbf{x}\rho$.

Case 3, $d_{\psi_1} \geqslant 1$, $d_{\psi_2} = 0$. The argument in this case is similar to that for case 2.

Case 4, $d_{\psi_1} \geqslant 1$ and $d_{\psi_2} \geqslant 1$. Then ψ'_1 and ψ'_2 exist and by the inductive hypothesis $\mathbf{x}_1\rho \in \hat{Q}'_{\psi'_1}(\gamma_{k_1}(G))$ and $\mathbf{x}_2\rho \in \hat{Q}'_{\psi'_2}(\gamma_{k_2}(G))$. But then, by theorem 3·1 (iv) and (v),

$$\mathbf{x}\rho = [\mathbf{x}_1\rho, \mathbf{x}_2\rho] \in \hat{Q}'_{\psi'_1+\psi'_2}(\gamma_{k_1}(G)) = \hat{Q}'_{\psi'}(\gamma_{k_1}(G)) \leqslant \hat{Q}'_{\alpha'}(\gamma_{k_1}(G)).$$

Now the converse inclusion will be proved, that $\hat{Q}'_{\phi'}(\gamma_{k_1}(G)) \leqslant \hat{Q}_\phi(G)$. By definitions 1·3 and 3·2, $\mathbf{N}^-_{k_1}$ is the set of all expressions of weight $\geqslant k_1$. It has just been remarked that this is a describable algebra, so there exists some description $\rho' : \mathbf{A}' \to \mathbf{N}^-_{k_1}$ where $\mathbf{A}'$ is some algebra of expressions: $\mathbf{A}'$ may be the same as $\mathbf{A}$ provided the latter has enough generators. Here two algebras of expressions are involved, so a little care must be taken with the definitions of $\hat{\mathbf{Q}}_\phi$ and $\hat{\mathbf{Q}}'_{\phi'}$ (see definition 3·2, where it was assumed that only one algebra of expressions was being considered). It will be assumed that $\hat{\mathbf{Q}}_\phi \subseteq \mathbf{A}$ and $\hat{\mathbf{Q}}'_{\phi'} \subseteq \mathbf{A}'$ or, more precisely,

$$\hat{\mathbf{Q}}_\phi = \{\mathbf{x} : \mathbf{x} \in \mathbf{A},\, \hat{\pi}(\mathbf{x}) \geqslant \phi\} \quad \text{and} \quad \hat{\mathbf{Q}}'_{\phi'} = \{\mathbf{x}' : \mathbf{x}' \in \mathbf{A}',\, \hat{\pi}'(\mathbf{x}') \geqslant \phi'\}.$$

Then $\hat{Q}_\phi(G) = \hat{\mathbf{Q}}_\phi\rho$ and $\hat{Q}_\phi(\gamma_{k_1}(G)) = \hat{\mathbf{Q}}'_{\phi'}\rho'\rho$. It is now sufficient to show that $\hat{\mathbf{Q}}'_{\phi'}\rho' \subseteq \hat{\mathbf{Q}}_\phi$. This is accomplished by proving that

$$\mathbf{x}' \in \hat{\mathbf{Q}}'_{\phi'} \Rightarrow \mathbf{x}'\rho' \in \hat{\mathbf{Q}}_\phi$$

by induction over the height of $\mathbf{x}$. If $\text{ht}(\mathbf{x}') = 1$ then either $\mathbf{x}' = \mathbf{1}'$ (the identity element of $\mathbf{A}'$), in which case $\mathbf{x}'\rho' = \mathbf{1} \in \hat{\mathbf{Q}}_\phi$, or else $\mathbf{x}' = \mathbf{g}'_i$ (one of the generators of $\mathbf{A}'$) in which case $\hat{\pi}'(\mathbf{x}') = 1$ so that $\phi(0) = k_1$ and $\phi(1) = 1$. Thus $\hat{\mathbf{Q}}_\phi = \mathbf{N}^-_{k_1}$ and so $\mathbf{x}'\rho' \in \mathbf{A}'\rho' = \mathbf{N}^-_{k_1} = \hat{\mathbf{Q}}_\phi$. Now suppose that $\text{ht}(\mathbf{x}) > 1$ and the result is true for all expressions of smaller height. If $\mathbf{x} = \mathbf{u}^{-1}$ or $\mathbf{x}_1\mathbf{x}_2$ the result follows immediately from the fact that $\hat{\mathbf{Q}}_\phi$ is a subalgebra. If $\mathbf{x}' = [\mathbf{x}'_1, \mathbf{x}'_2]$ write $\hat{\pi}'(\mathbf{x}') = \psi'$ (so that $\psi' \geqslant \phi'$), $\hat{\pi}'(\mathbf{x}'_1) = \psi'_1$ and $\hat{\pi}'(\mathbf{x}'_2) = \psi'_2$. Then $\psi' = \psi'_1 + \psi'_2$. Define $\psi \in Q$ by $\psi(j) = \psi'(j-1)$ for $j \geqslant 1$ and $\psi(0) = k_1\psi(1)$ and define $\psi_1, \psi_2 \in Q$ similarly. Then $\psi \geqslant \phi$ and, by the inductive hypothesis,

$$\mathbf{x}_1\rho' \in \hat{Q}_{\psi_1} \quad \text{and} \quad \mathbf{x}_2\rho' \in \hat{Q}_{\psi_2}$$

so that $\mathbf{x}'\rho \in \hat{\mathbf{Q}}_{\psi_1+\psi_2} = \hat{\mathbf{Q}}_\psi \subseteq \hat{\mathbf{Q}}_\phi$.

Corollary 1. *With the notation and premises of the lemma,*

$$\hat{\mathfrak{Q}}_\phi = \hat{\mathfrak{Q}}'_{\phi'} . \mathfrak{N}_{k_1-1}.$$

Corollary 2. *With the notation of the lemma, suppose* $\Phi \subseteq Q$ *has the property that, for any* $\phi \in \Phi$, $\phi(0) = k_1\phi(1)$. *Then*

$$\hat{\mathfrak{Q}}_\Phi = \hat{\mathfrak{Q}}'_{\Phi'} . \mathfrak{N}_{k_1-1}$$

where $\Phi' = \{\phi' : \phi \in \Phi\}$.

Proof. This follows from the lemma and theorem 3·1 (viii).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Corollary 3. *With the notation and premises of the lemma,*

$$\mathfrak{Q}_\phi = \mathfrak{Q}'_{\phi'} . \mathfrak{N}_{k_1-1}.$$

Proof. Since a constructive definition has been given for polyweights which precludes the possibility of the fine and coarse orders being the same, this result may not be inferred from corollary 1 by means of the metatheorem of § 3.

However, since $\phi(0) = k_1 \phi(1)$, lemma 11·2 gives

$$\begin{aligned} Q_\phi(G) &= \prod_{\substack{\xi \geq \phi \\ \xi(0)=k_1\xi(1)}} \hat{Q}_\xi(G) \\ &= \prod_{\xi' \geq \phi'} \hat{Q}'_{\xi'}(\gamma_{k_1}(G)) \\ &= Q'_{\xi'}(\gamma_{k_1}(G)) \end{aligned}$$

for any group G.

Theorem 15·1. Any polyweight range is partially collectable.

Proof. Write Q^S for the polyweight range defined by the sequence $S = (2, 2, 2, \ldots)$. From definition 14·2 it follows immediately that every polyweight range Q^K is a subset of Q^S and the coarse order defined on Q^K is the restriction to Q^K of that defined on Q^S and thus, if Φ and Ψ are two subsets of Q^K then $\Phi \preccurlyeq \Psi$ in Q^K if and only if $\Phi \preccurlyeq \Psi$ in Q^S and Φ is a totally unordered subset of Q^K if and only if it is a totally unordered subset of Q^S. This makes possible the comparison of subsets of distinct polyweight ranges in a consistent manner.

The theorem is proved using lemma 13·1 (iii) by showing that the relatively free group $F(\hat{\mathfrak{Q}}^K_\Phi)$ of any rank is residually a finite p-group for any prime p: notice that $F(\hat{\mathfrak{Q}}^K_\Phi) = F(\hat{\mathfrak{Q}}^K_{\Phi_0})$ where Φ_0 is the set of minimal elements in Φ, so it is sufficient to prove this statement only for totally unordered sets Φ. This is accomplished by induction over the totally unordered subsets of Q^S using part (ii) of the corollary to lemma 4·1 so the inductive assumption becomes: if Ψ is a totally unordered subset of any polyweight range Q^L and $\Psi \prec \Phi$ in Q^S then the relatively free group $F(\hat{\mathfrak{Q}}^L_\Psi)$ of any rank is residually a finite p-group for any prime p.

Assume then that Φ is a totally unordered subset of Q^K and the inductive hypothesis is true. Write $G = F(\hat{\mathfrak{Q}}^K_\Phi)$. From the definition of addition there are four possibilities: $1 \in \Phi$, $\infty \in \Phi$, there exists $\psi \in Q^K$ such that $\psi + 1 \in \Phi$ or, for every $\phi \in \Phi$, $\phi(0) = k_1 \phi(1)$.

If $1 \in \Phi$ then, since Φ is totally unordered, $\Phi = \{1\}$ so G is trivial and is thus (vacuously) residually a finite p-group.

If $\infty \in \Phi$ then $\Phi = \{\infty\}$ so G is an absolutely free group which is residually a finite p-group.

If there exists $\psi \in Q^K$ such that $\psi + 1 \in \Phi$, write Ψ for the subset of Q^K made by replacing $\psi + 1$ by ψ in Φ, that is, $\Psi = (\Phi - \{\psi + 1\}) \cup \{\psi\}$ and Ψ_0 for the set of minimal elements in Ψ. Then $\Psi_0 \prec \Phi$ so $F(\hat{\mathfrak{Q}}^K_{\Psi_0})$ is residually a finite p-group. Now let $x \in G - \{1\}$. Then either $x \notin \hat{\mathfrak{Q}}^K_{\Psi_0}(G)$, in which case there exists a normal subgroup N of G such that $x \notin N$ and G/N is a finite p-group since $G/\hat{Q}^K_{\Psi_0}(G) \cong F(\hat{\mathfrak{Q}}^K_{\Psi_0})$, or else $x \in \hat{Q}^K_{\Psi_0}(G)$. But Ψ_0 consists of ψ together with a subset of Φ so Φ does not coarsely dominate Ψ_0: then by the basis theorem, theorem 9·1 (C) (iii), there exists a non-trivial Q^K-basic expression $\mathbf{x} \in \hat{\mathbf{B}}^{Q^K}_\Phi \cap \hat{\mathbf{Q}}^K_{\Psi_0}$ such that $\mathbf{x}\rho = x$. But then there exists a non-negative integer c such that $\mathbf{x} \in \mathbf{B}^{Q^K}_{(c)}$ so $x \notin \gamma_c(G)$ by theorem 9·1 (D) (iii) which also implies that $G/\gamma_c(G)$ is torsion free and hence residually a finite p-group.

Finally, suppose $\phi(0) = k_1\phi(1)$ for every $\phi \in \Phi$. Define K' and Φ' as in lemma 15·1, corollary 2. By that corollary $\hat{\mathfrak{Q}}^K_\Phi = \hat{\mathfrak{Q}}^{K'}_{\Phi'}.\mathfrak{N}_{k_1-1}$ and by the inductive hypothesis $F(\hat{\mathfrak{Q}}^{K'}_{\Phi'})$ is residually a finite p-group. But so is $F(\mathfrak{N}_{k_1-1})$ and therefore, by a theorem of Baumslag (1963, theorem 3) G is residually a finite p-group.

16. *Polyweight subgroups as products of commutator subgroups*

THEOREM 16·1. *Let $K = (k_i)_{i=1}^{\infty}$ be a sequence of integers, each $\geqslant 2$, let $Q = Q^K$ be the corresponding polyweight range. Then for each non-negative integer r,*

(i) $\hat{\mathfrak{Q}}_{\delta_r} = \mathfrak{Q}_{\delta_r} = \mathfrak{P}_{K_r}$ *and*

(ii) *for any group G, $\hat{Q}_{\delta_r}(G) = Q_{\delta_r}(G) = P_{K_r}(G)$,*

where $\delta_r = \delta_r^K$ is the function given in definition 14·3.

Proof. (i) follows immediately from lemma 15·1 and its third corollary by induction over r and (ii) is a corollary of (i).

The major part of this section will be occupied by a closer study of the formula of lemma 5·1 for polyweights: first a constructive definition of $\hat{U}(\phi)$ is given and then some of the redundancy removed from that formula, resulting in practical recursive expressions for the polyweight subgroups of an arbitrary group G.

For the remainder of this section it will be assumed that a fixed polyweight range $Q = Q^K$ is being considered.

DEFINITION 16·1

(i) *For each function $\phi \in Q$ and each non-negative integer h, the function $\phi_{+h}: \omega \to \omega$ is defined: if $h = d_\phi$ and $\phi(h) = k_{h+1} - 1$ then*

$$\begin{aligned} \phi_{+h}(j) &= \phi(j) && (j \neq h \quad \textit{or} \quad h+1) \\ &= \phi(h)+1 && (j = h) \\ &= 1 && (j = h+1), \end{aligned}$$

and otherwise

$$\begin{aligned} \phi_{+h}(j) &= \phi(j) && (j \neq h) \\ &= \phi(h)+1 && (j = h). \end{aligned}$$

The function $\phi_{-h}: \omega \to \omega \cup \{-1\}$ is defined

$$\begin{aligned} \phi_{-h}(j) &= \phi(j) && (j \neq h) \\ &= \phi(h)-1 && (j = h). \end{aligned}$$

The functions ϕ_{+h} and ϕ_{-h} are not necessarily members of Q, for instance 1_{+1} and 1_{-1} are not.

(ii) *The sets $H^+(\phi)$ and $H^-(\phi)$ are defined to be the set of positive integers h such that $\phi_{+h} \in Q$ and the set of non-negative integers such that $\phi_{-h} \in Q$ respectively, that is,*

$$H^+(\phi) = \{h: h \geqslant 1, \phi(h-1) \geqslant k_h(\phi(h)+1)\},$$

$$H^-(\phi) = \{h: \phi(h) > k_{h+1}\phi(h+1), \phi(h) > 1\}.$$

(iii) *The sets $\hat{U}_1(\phi)$ and $\hat{U}_2(\phi)$ are defined*

$$\hat{U}_1(\phi) = \{(\alpha, \beta): \alpha \geqslant \beta, \alpha+\beta = \phi\},$$

$$\hat{U}_2(\phi) = \{(\alpha,\beta): \alpha+\beta = \phi_{+h} \quad \textit{where} \quad h \in H^+(\phi) - H^-(\phi), \alpha \geqslant \beta, \beta(h) = 1\}.$$

LEMMA 16·1. $\hat{U}(\phi) = \hat{U}_1(\phi) \cup \hat{U}_2(\phi)$.

Proof. First it is shown that $\hat{U}(\phi) \subseteq \hat{U}_1(\phi) \cup \hat{U}_2(\phi)$. It is sufficient to show that

$$\hat{U}(\phi) - \hat{U}_1(\phi) \subseteq \hat{U}_2(\phi).$$

Suppose then that $(\alpha, \beta) \in \hat{U}(\phi) - \hat{U}_1(\phi)$. Then $\alpha+\beta \succ \phi$ so $(\alpha+\beta)(j) \geqslant \phi(j)$ for all $j \in \omega$ and there exists $h \in \omega$ such that $(\alpha+\beta)(h) > \phi(j)$: suppose that h is the largest such integer. Then $d_{\alpha+\beta} \geqslant h$. Now $\alpha \geqslant \beta$ so $d_\alpha \geqslant d_\beta$ and five cases arise: case 1, $d_\alpha = d_\beta = h$ and $\alpha(h)+\beta(h) < k_{h+1}$ so $d_{\alpha+\beta} = h$, case 2, $d_\alpha = d_{\alpha+\beta} = h > d_\beta$, case 3, $d_\alpha = d_\beta = h-1$ and $\alpha(h-1)+\beta(h-1) \geqslant k_h$ so $d_{\alpha+\beta} = h$, case 4, $d_\alpha = d_\beta = h$ and $\alpha(h)+\beta(h) \geqslant k_{h+1}$ so $d_{\alpha+\beta} = h+1$ and case 5, $d_\alpha > h$.

Case 1, $d_\alpha = d_\beta = h$, $\alpha(h)+\beta(h) < k_{h+1}$ and $d_{\alpha+\beta} = h$. Suppose $\alpha(h) > 1$. Then $h \in H^-(\alpha)$ so $\alpha_{-h} \in Q$ and then $\alpha_{-h} \prec \alpha$ and $\alpha_{-h}+\beta \succcurlyeq \phi$ which contradicts the assumption that $(\alpha, \beta) \in \hat{U}(\phi)$. Thus $\alpha(h) = 1$ and by a similar argument $\beta(h) = 1$. Thus $(\alpha+\beta)(h) = 2$ so $\phi(h) = 0$ or 1. This gives rise to three subcases: case 1·1, $\phi(h) = 0$, case 1·2, $\phi(h) = 1$ and $\phi(h-1) < 2k_h$ and case 1·3, $\phi(h) = 1$ and $\phi(h-1) \geqslant 2k_h$.

Case 1·1, $\phi(h) = 0$. Define α' by

$$\begin{aligned} \alpha'(j) &= \alpha(j) \quad && (j < h-1) \\ &= 1 && (j = h-1) \\ &= 0 && (j > h-1). \end{aligned}$$

Then $\alpha' \in Q$ and $\alpha' \prec \alpha$. Further

$$(\alpha'+\beta)(j) = (\alpha+\beta)(j) \geqslant \phi(j) \quad (j < h-1),$$
$$(\alpha'+\beta)(h-1) \geqslant \beta(h-1) \geqslant k_h > \phi(h-1)$$

since $d_\beta = h$ and $\phi(h) = 0$. Thus $\alpha'+\beta \succcurlyeq \phi$ contradicting $(\alpha, \beta) \in \hat{U}(\phi)$, so $\phi(h) = 0$ is impossible.

Case 1·2, $\phi(h) = 1$ and $\phi(h-1) < 2k_h$. Define β' by

$$\begin{aligned} \beta'(j) &= \beta(j) \quad && (j < h-1), \\ &= k_h - 1 && (j = h-1), \\ &= 0 && (j > h-1). \end{aligned}$$

Then $\beta' \in Q$ and $\beta' \prec \beta$. Further

$$(\alpha+\beta')(j) = (\alpha+\beta)(j) \geqslant \phi(j) \quad (j < h-1),$$
$$(\alpha+\beta')(h-1) = \alpha(h-1)+k_h-1 \geqslant 2k_h-1 \geqslant \phi(h-1),$$
$$(\alpha+\beta')(h) = \alpha(h) = 1 = \phi(h),$$

so $\alpha+\beta' \succcurlyeq \phi$ which is another contradiction, so this subcase is also impossible.

Case 1·3, $\phi(h) = 1$ and $\phi(h-1) \geqslant 2k_h$. Then

$$\phi(h-1) \geqslant k_h(\phi(h)+1) \quad \text{and} \quad \phi(h) = 1 \quad \text{so} \quad h \in H^+(\phi) - H^-(\phi).$$

It has already been observed that $\beta(h) = 1$, $(\alpha+\beta)(h) = 2 = \phi(h)+1$ and, for $j > h$, $(\alpha+\beta)(j) = 0 = \phi(j)$. It thus remains to show that $(\alpha+\beta)(j) = \phi(j)$ for $j < h$. Suppose $(\alpha+\beta)(h-1) > \phi(h-1)$. Then $\alpha(h-1)+\beta(h-1) > \phi(h-1) \geqslant 2k_h$ and so either $\alpha(h-1) > k_h$ or $\beta(h-1) > k_h$. If $\alpha(h-1) > k_h$ then $h-1 \in H^-(\alpha)$ and then $\alpha_{-(h-1)} \in Q$, $\alpha_{-(h-1)} \prec \alpha$ and $\alpha_{-(h-1)}+\beta \succcurlyeq \phi$ which contradicts $(\alpha, \beta) \in \hat{U}(\phi)$. Similarly, $\beta(h-1) > k_h$ yields a contradiction, so $(\alpha+\beta)(h-1) = \phi(h-1)$.

Now suppose, for some $j < h-1$, $(\alpha+\beta)(j) > \phi(j)$. Let j be the largest such integer, so that $\alpha(j)+\beta(j) > \phi(j)$ and $\alpha(j+1)+\beta(j+1) = \phi(j+1)$. Then either $\alpha(j) > k_{j+1}\alpha(j+1)$ or $\beta(j) > k_{j+1}\beta(j+1)$. If $\alpha(j) > k_{j+1}\alpha(j+1)$ then $j \in H^-(\alpha)$ so $\alpha_{-j} \in Q$, $\alpha_{-j} \prec \alpha$ and $\alpha_{-j}+\beta \succcurlyeq \phi$ contradicting $(\alpha,\beta) \in \hat{U}(\phi)$. Similarly, $\beta(j) > k_{j+1}\beta(j+1)$ yields a contradiction so $(\alpha+\beta)(j) = \phi(j)$ for all $j < h-1$.

Case 2, $d_\alpha = h > \beta$. If $\alpha(h) > 1$ then $h \in H^-(\alpha)$ so $\alpha_{-h} \in Q$, $\alpha_{-h} \prec \alpha$ and $\alpha_{-h}+\beta \succcurlyeq \phi$ which is a contradiction. Thus $\alpha(h) = (\alpha+\beta)(h) = 1$ so $\phi(h) = 0$ and $d_\phi < h$. Define α' by

$$\begin{aligned} \alpha'(j) &= \alpha(j) && (j < h-1) \\ &= k_h - 1 && (j = h-1) \\ &= 0 && (j > h-1). \end{aligned}$$

Then $\alpha' \in Q$, $\alpha' \prec \alpha$ and $\alpha'+\beta \succcurlyeq \phi$ which contradicts $(\alpha,\beta) \in \hat{U}(\phi)$ so this case is impossible.

Case 3, $d_\alpha = d_\beta = h-1$, $\alpha(h-1)+\beta(h-1) \geqslant k_h$ and $d_{\alpha+\beta} = h$. Then $(\alpha+\beta)(h) = 1$ so $\phi(h) = 0$ and $d_\phi < h$. Suppose that $\alpha(h-1) \geqslant 2$. Then $h-1 \in H^-(\alpha)$ so $\alpha_{-(h-1)} \in Q$, $\alpha_{-(h-1)} \prec \alpha$ and $\alpha_{-(h-1)}+\beta \succcurlyeq \phi$ which is a contradiction. Thus $\alpha(h-1) = 1$ and by a similar argument $\beta(h-1) = 1$. Then $k_h = 2$; but $\phi(h) = 0$ so $\phi(h-1) = 0$ or 1.

Suppose that either $\phi(h-1) = 0$ or $\phi(h-2) < 2k_{h-1}$. Define α' by

$$\begin{aligned} \alpha'(j) &= \alpha(j) && (j < h-2), \\ &= k_{h-1} - 1 && (j = h-2), \\ &= 0 && (j > h-2). \end{aligned}$$

Then
$$\begin{aligned} (\alpha'+\beta)(j) &= (\alpha+\beta)(j) \geqslant \phi(j) && (j < h-2), \\ (\alpha'+\beta)(h-2) &\geqslant 2k_{h-1} - 1 \geqslant \phi(h-2) && (\text{since } d_\beta = h-1), \\ (\alpha'+\beta)(h-1) &= \beta(h-1) \geqslant 1 = \phi(h-1), \end{aligned}$$

so $\alpha' \in Q$, $\alpha' \prec \alpha$ and $\alpha'+\beta \succcurlyeq \phi$ which is a contradiction. Thus $\phi(h-1) = 1$ and $\phi(h-2) \geqslant 2k_{h-1}$ so $h-1 \in H^+(\phi) - H^-(\phi)$. Since $k_h = 2$, $\phi_{+(h-1)}$ is given by

$$\begin{aligned} \phi_{+(h-1)}(j) &= \phi(j) && (j < h-1) \\ &= \phi(h-1)+1 = 2 && (j = h-1) \\ &= 1 && (j = h) \\ &= 0 && (j > h) \end{aligned}$$

so that $(\alpha+\beta)(j) = \phi_{+(h-1)}(j)$ for $j \geqslant h-1$. Since $\beta(h) = 1$, it remains to show that $(\alpha+\beta)(j) = \phi(j)$ for $j < h-1$: the argument to this end is the same as that given in case 1·3.

Case 4, $d_\alpha = d_\beta = h$, $\alpha(h)+\beta(h) \geqslant k_{h+1}$ and $d_{\alpha+\beta} = h+1$. Then $(\alpha+\beta)(h+1) = 1$ and, by the choice of h, $\phi(h+1) = 1$. Thus

$$a(h)+\beta(h) = (\alpha+\beta)(h) > \phi(h) \geqslant k_{h+1} \geqslant 2$$

and, since $\alpha \geqslant \beta$, $\alpha(h) \geqslant \beta(h)$ so $\alpha(h) \geqslant 2$. Thus $h \in H^-(\alpha)$ which yields a contradiction as usual and this case is impossible.

Case 5, $d_\alpha > h$. Then $(\alpha+\beta)(j) = \alpha(j)+\beta(j)$ for $j \leqslant h+1$ so

$$\alpha(h)+\beta(h) > \phi(h) \geqslant k_{h+1}\phi(h+1) = k_{h+1}\alpha(h+1)+k_{h+1}\beta(h+1)$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

by the choice of h and then either

$$\alpha(h) > k_{h+1}\alpha(h+1) \quad \text{or} \quad \beta(h) > k_{h+1}\beta(h+1).$$

If $\alpha(h) > k_{h+1}\alpha(h+1)$ then, since $d_\alpha > h$, $\alpha(h) > 1$ so $h \in H^-(\alpha)$ which yields a contradiction as usual. Thus $\alpha(h) = k_{h+1}\alpha(h+1)$ and $\beta(h) > k_{h+1}\beta(h+1)$. If $\beta(h) > 1$ the same argument yields a contradiction so $\beta(h) = 1$.

Now $$\phi(h) < \alpha(h)+\beta(h) = \alpha(h)+1 = k_{h+1}\alpha(h+1)+1 = k_{h+1}\phi(h+1)+1$$

so $\phi(h) = k_{h+1}\phi(h+1)$, $h \notin H^-(\phi)$, $\alpha(h) = \phi(h)$ and $(\alpha+\beta)(h) = \phi(h)+1$.

Suppose $\phi(h-1) < k_h(\phi(h)+1)$. Define β' by

$$\begin{aligned} \beta'(j) &= \beta(j) && (j < h-1),\\ &= k_h - 1 && (j = h-1),\\ &= 0 && (j > h-1). \end{aligned}$$

Then $\beta' \in Q$, $\beta' \prec \beta$ and $\quad (\alpha+\beta')(j) = (\alpha+\beta)(j) \quad (j < h-1)$,

$$(\alpha+\beta')(h-1) = \alpha(h-1)+k_h-1 \geqslant k_h(\alpha(h)+1)-1 = k_h(\phi(h)+1) \geqslant \phi(h-1),$$

$$(\alpha+\beta')(j) = \alpha(j) = \phi(j) \quad (j > h-1),$$

so $\alpha+\beta' \geqslant \phi$ which is a contradiction. Thus $\phi(h-1) \geqslant k_h(\phi(h)+1)$ and $h \in H^+(\phi)$. Since $\beta(h) = 1$ and $\phi_{+h}(j) = (\alpha+\beta)(j)$ for $j > h-1$ it remains to show that $(\alpha+\beta)(j) = \phi(j)$ for $j \leqslant h-1$.

Suppose $(\alpha+\beta)(h-1) > \phi(h-1)$. Then

$$\alpha(h-1)+\beta(h-1) > \phi(h-1) \geqslant k_h(\phi(h)+1) = k_h(\alpha(h)+\beta(h))$$

so either $\alpha(h-1) > k_h\alpha(h)$ or $\beta(h-1) > k_h\beta(h)$ both of which possibilities yield contradictions as usual. Thus $(\alpha+\beta)(h-1) = \phi(h-1)$. Suppose there exists $j < h-1$ such that $(\alpha+\beta)(j) > \phi(j)$. Let j be the largest such integer. Then

$$\alpha(j)+\beta(j) > \phi(j) \quad \text{and} \quad \alpha(j+1)+\beta(j+1) = \phi(j+1)$$

so either $\alpha(j) > k_{j+1}\alpha(j+1)$ or $\beta(j) > k_{j+1}\beta(j+1)$ both of which possibilities again yield contradictions.

This completes the proof that $\hat{U}(\phi) \subseteq \hat{U}_1(\phi) \cup \hat{U}_2(\phi)$. That $\hat{U}_1(\phi) \subseteq \hat{U}(\phi)$ follows immediately from the definitions so it remains to show that $\hat{U}_2(\phi) \subseteq \hat{U}(\phi)$.

Suppose then that $(\alpha,\beta) \in \hat{U}_2(\phi)$, $\alpha+\beta = \phi_{+h}$. Then $\alpha+\beta \geqslant \phi$ immediately. Since $0 \notin H^+(\phi)$, $(\alpha+\beta)(0) = \alpha(0)+\beta(0) = \phi(0)$ so $\alpha(0) < \phi(0)$ and $\beta(0) < \phi(0)$: thus $\alpha \ngeqslant \phi$ and $\beta \ngeqslant \phi$. By the definition of $\hat{U}_2(\phi)$, $\alpha \geqslant \beta$. Now suppose $\alpha' \prec \alpha$. Then there exists j such that $\alpha'(j) < \alpha(j)$. Now $\beta(h) = 1$ so $\alpha(h)+1 = \phi_{+h}(h) = \phi(h)+1$ and then $\alpha(h) = \phi(h)$. Also, since $d_\beta = h$, $\alpha(h+1) = \phi(h+1)$ so, since $h \notin H^-(\phi)$, either

$$\alpha(h) = \phi(h) = 1 \quad \text{or} \quad \alpha(h) = \phi(h) = k_{h+1}\phi(h+1) = k_{h+1}\alpha(h+1).$$

In either case $\alpha'(h) < \alpha(h) \Rightarrow \alpha'(h-1) \leqslant \alpha(h-1)$ so without loss of generality it may be assumed that $j < h$. Then

$$(\alpha'+\beta)(j) = \alpha'(j)+\beta(j) < \alpha(j)+\beta(j) = \phi(j) \quad \text{and} \quad \alpha'+\beta \ngeqslant \phi.$$

By a similar argument, $\beta' \prec \beta \Rightarrow \alpha+\beta' \ngeqslant \phi$.

Lemma 16·2. *Suppose $\phi \in Q$, $\phi \neq 1$ and G is any group. Then*

$$\hat{Q}_\phi(G) = \prod_{(\alpha,\beta)\in\hat{U}_1(\phi)} [\hat{Q}_\alpha(G), \hat{Q}_\beta(G)].$$

Proof. The argument proceeds by induction over ϕ. By virtue of lemmas 5·1 and 16·1 it is sufficient to show that, for each $(\alpha,\beta)\in\hat{U}_2(\phi)$,

$$[\hat{Q}_\alpha(G), \hat{Q}_\beta(G)] \leqslant \prod_{(\xi,\zeta)\in\hat{U}_1(\phi)} [\hat{Q}_\xi(G), \hat{Q}_\zeta(G)].$$

This is proved by a subsidiary induction over β.

By the main inductive hypothesis,

$$\hat{Q}_\beta(G) = \prod_{(\beta_1,\beta_2)\in\hat{U}_1(\beta)} [\hat{Q}_{\beta_1}(G), \hat{Q}_{\beta_2}(G)]$$

and consequently, since all these subgroups are normal,

$$[\hat{Q}_\alpha(G), \hat{Q}_\beta(G)] = \prod_{(\beta_1,\beta_2)\in\hat{U}_1(\beta)} [\hat{Q}_\alpha(G), [\hat{Q}_{\beta_1}(G), \hat{Q}_{\beta_2}(G)]].$$

It now becomes sufficient to show that, if $(\alpha,\beta)\in\hat{U}_2(\phi)$ and $(\beta_1,\beta_2)\in\hat{U}_1(\beta)$ then

$$[\hat{Q}_\alpha(G), [\hat{Q}_{\beta_1}(G), \hat{Q}_{\beta_2}(G)]] \leqslant \prod_{(\xi,\zeta)\in\hat{U}_1(\phi)} [\hat{Q}_\xi(G), \hat{Q}_\zeta(G)].$$

Since $(\alpha,\beta)\in\hat{U}_2(\phi)$, there exists $h\in H^+(\phi)$ such that $\alpha+\beta=\phi_{+h}$ and $\beta(h)=1$. Since

$$(\beta_1,\beta_2)\in\hat{U}_1(\beta), \quad \beta_1+\beta_2=\beta \quad \text{and} \quad \beta_1 \geqslant \beta_2.$$

This gives rise to two cases: case 1, $d_{\beta_1}=d_{\beta_2}=h-1$ and $\beta_1(h-1)+\beta_2(h-1)\geqslant k_h$ and case 2, $d_{\beta_1}=h>d_{\beta_2}$ and $\beta_1(h)=1$.

Case 1, $d_{\beta_1}=d_{\beta_2}=h-1$ and $\beta_1(h-1)+\beta_2(h-1)\geqslant k_h$. Then $\beta=\beta_1+\beta_2$ is given by

$$\begin{aligned}\beta(j) &= \beta_1(j)+\beta_2(j) && (j<h)\\ &= 1 && (j=h)\\ &= 0 && (j>h).\end{aligned}$$

There are two possibilities for the function $\alpha+\beta$. Since $\alpha\geqslant\beta$, $d_\alpha\geqslant h$. If

$$d_\alpha = h \quad \text{and} \quad \alpha(h)=k_{h+1}-1$$

then $\alpha+\beta$ is given by

$$\begin{aligned}(\alpha+\beta)(j) &= \alpha(j)+\beta_1(j)+\beta_2(j) && (j<h)\\ &= k_{h+1} && (j=h)\\ &= 1 && (j=h+1)\\ &= 0 && (j>h+1),\end{aligned}$$

and otherwise

$$\begin{aligned}(\alpha+\beta)(j) &= \alpha(j)+\beta_1(j)+\beta_2(j) && (j<h)\\ &= \alpha(j)+1 && (j=h)\\ &= \alpha(j) && (j>h).\end{aligned}$$

But $\alpha+\beta=\phi_{+h}$ so in either case ϕ is given by

$$\phi(j)=\alpha(j)+\beta_1(j)+\beta_2(j) \quad (\text{for all } j).$$

Now consider $\alpha+\beta_1$. Since $d_\alpha \geqslant h$ and $d_{\beta_1} = h-1$,

$$(\alpha+\beta_1)(j) = \alpha(j)+\beta_1(j) \quad \text{(for all } j\text{)},$$

and then $d_{\alpha+\beta_1} = d_\alpha \geqslant h$ and $d_{\beta_1} = h-1$ so

$$(\alpha+\beta_1+\beta_2)(j) = \alpha(j)+\beta_1(j)+\beta_2(j) \quad \text{(for all } j\text{)},$$

and thus $\alpha+\beta_1+\beta_2 = \phi$. But $\alpha \geqslant \beta > \beta_1$ so $\alpha \geqslant \beta_1 \geqslant \beta_2$ and then $\alpha+\beta_2+\beta_1 = \alpha+\beta_1+\beta_2 = \phi$. Thus $(\alpha+\beta_1, \beta_2)$ and $(\alpha+\beta_2, \beta_1)$ are both members of $\hat{U}_1(\phi)$. But

$$\begin{aligned}[\hat{Q}_\alpha(G), [\hat{Q}_{\beta_1}(G), \hat{Q}_{\beta_2}(G)]] &= [\hat{Q}_{\beta_1}(G), \hat{Q}_{\beta_2}(G), \hat{Q}_\alpha(G)] \\ &\leqslant [\hat{Q}_{\beta_2}(G), \hat{Q}_\alpha(G), \hat{Q}_{\beta_1}(G)]\,[\hat{Q}_\alpha(G), \hat{Q}_{\beta_1}(G), \hat{Q}_{\beta_2}(G)], \\ &\leqslant [\hat{Q}_{\beta_2+\alpha}(G), \hat{Q}_{\beta_1}(G)]\,[\hat{Q}_{\alpha+\beta_1}(G), \hat{Q}_{\beta_2}(G)]\end{aligned}$$

by theorem 3·1 (v) which completes the proof for case 1.

Case 2, $d_{\beta_1} = h > d_{\beta_2}$ and $\beta_1(h) = 1$. Two subcases are now considered: case 2·1, $d_\alpha = h$ and $\alpha(h) = k_{h+1}-1$ and case 2·2, either $d_\alpha \neq h$ or $\alpha(h) \neq k_{h+1}-1$.

Case 2·1, $d_\alpha = h$ and $\alpha(h) = k_{h+1}-1$. Then

$$\beta(j) = \beta_1(j)+\beta_2(j) \quad \text{(for all } j\text{)},$$

so

$$\begin{aligned}(\alpha+\beta)(j) &= \alpha(j)+\beta_1(j)+\beta_2(j) && (j<h) \\ &= k_{h+1} && (j=h) \\ &= 1 && (j=h+1) \\ &= 0 && (j>h+1),\end{aligned}$$

and $\alpha+\beta = \phi_{+h}$ so ϕ is given by

$$\begin{aligned}\phi(j) &= \alpha(j)+\beta_1(j)+\beta_2(j) && (j<h) \\ &= k_{h+1}-1 && (j=h) \\ &= 0 && (j>h).\end{aligned}$$

Now $d_\alpha = h$, $\alpha(h) = k_{h+1}-1$ and $\beta_1(h) = 1$ so

$$\begin{aligned}(\alpha+\beta_1)(j) &= \alpha(j)+\beta_1(j) && (j<h) \\ &= k_{h+1} && (j=h) \\ &= 1 && (j=h+1) \\ &= 0 && (j>h+1).\end{aligned}$$

Define a function ψ by

$$\begin{aligned}\psi(j) &= \alpha(j)+\beta_1(j) && (j<h) \\ &= k_{h+1}-1 && (j=h) \\ &= 0 && (j>h).\end{aligned}$$

Then $\psi \geqslant \alpha+\beta_1$ so $\hat{Q}_{\alpha+\beta_1}(G) \leqslant \hat{Q}_\psi(G)$. But $d_\psi = h$ and $d_{\beta_2} < h$ so $\psi+\beta_2$ is given by

$$\begin{aligned}(\psi+\beta_2)(j) &= \alpha(j)+\beta_1(j)+\beta_2(j) && (j<h) \\ &= k_{h+1}-1 && (j=h) \\ &= 0 && (j>h),\end{aligned}$$

that is, $\psi+\beta_2 = \phi$. But $\psi \geqslant \beta_2$ so $(\psi, \beta_2) \in \hat{U}_1(\phi)$. Thus

$$[\hat{Q}_{\alpha+\beta_1}(G), \hat{Q}_{\beta_2}(G)] \leqslant [\hat{Q}_{\psi}(G), \hat{Q}_{\beta_2}(G)] \leqslant \prod_{(\xi,\zeta)\in\hat{U}_1(\phi)} [\hat{Q}_{\xi}(G), \hat{Q}_{\zeta}(G)].$$

Also $\alpha \geqslant \beta_1 \geqslant \beta_2$ so $\alpha+\beta_2+\beta_1 = \alpha+\beta_1+\beta_2 = \phi_{+h}$ and $\beta_1 < \beta$ so by the subsidiary inductive hypothesis,

$$[\hat{Q}_{\alpha+\beta_2}(G). \hat{Q}_{\beta_1}(G)] \leqslant \prod_{(\xi,\zeta)\in\hat{U}_1(\phi)} [\hat{Q}_{\xi}(G), \hat{Q}_{\zeta}(G)].$$

But $$[\hat{Q}_{\alpha}(G), [\hat{Q}_{\beta_1}(G), \hat{Q}_{\beta_2}(G)]] \leqslant [\hat{Q}_{\alpha+\beta_1}(G), \hat{Q}_{\beta_2}(G)]\,[\hat{Q}_{\alpha+\beta_2}(G), \hat{Q}_{\beta_1}(G)]$$

as before, which completes the proof in this case.

Case 2·2, either $d_\alpha \neq h$ or $\alpha(h) \neq k_{h+1}-1$. Then

$$(\alpha+\beta_1+\beta_2)\,(j) = (\alpha+\beta_2+\beta_1)\,(j) = (\beta_1+\beta_2+\alpha)\,(j) = \phi_{+h}(j) = \alpha(j)+\beta_1(j)+\beta_2(j)$$

for all j. The argument is now a simplified form of the one just given.

THEOREM 16·2. *Let Q be a polyweight range, G be any group and $\phi \in Q$, $\phi \neq 1$. Then, writing $d = d_\phi$,*

(i) *if $\phi(d) > 1$,* $$\hat{Q}_{\phi}(G) = \prod_{\substack{\alpha+\beta=\phi\\ \alpha\geqslant\beta\\ d_\beta=d}} [\hat{Q}_{\alpha}(G), \hat{Q}_{\beta}(G)]$$

(*notice that $d_\alpha = d$ also since $\beta \leqslant \alpha < \phi$*).

(ii) *If $\phi(d) = 1$,* $$\hat{Q}_{\phi}(G) = \prod_{\substack{\alpha+\beta=\phi\\ \alpha\geqslant\beta\\ d_\beta=d-1}} [\hat{Q}_{\alpha}(G), \hat{Q}_{\beta}(G)]$$

(*and here $d_\alpha = d$ or $d-1$*).

Proof. The argument proceeds by induction over ϕ.

(i) Using definition 16·1 (iii), lemma 16·2 becomes

$$\hat{Q}_{\phi}(G) = \prod_{\substack{\alpha+\beta=\phi\\ \alpha\geqslant\beta}} [\hat{Q}_{\alpha}(G), \hat{Q}_{\beta}(G)],$$

so it is sufficient to show that, if $\alpha+\beta = \phi$, $\alpha \geqslant \beta$ and $d_\beta < d$ then

$$[\hat{Q}_{\alpha}(G), \hat{Q}_{\beta}(G)] \leqslant \prod_{\substack{\xi+\zeta=\phi\\ \xi\geqslant\zeta\\ d_\xi=d}} [\hat{Q}_{\xi}(G), \hat{Q}_{\zeta}(G)].$$

Since $\phi(d) > 1$, $d_\alpha = d$ and $\alpha(d) = \phi(d) > 1$ so by the inductive hypothesis

$$\hat{Q}_{\alpha}(G) = \prod_{\substack{\alpha_1+\alpha_2=\alpha\\ \alpha_1\geqslant\alpha_2\\ d_{\alpha_1}=d}} [\hat{Q}_{\alpha_1}(G), \hat{Q}_{\alpha_2}(G)],$$

and so $$[\hat{Q}_{\alpha}(G), \hat{Q}_{\beta}(G)] = \prod_{\substack{\alpha_1+\alpha_2=\alpha\\ \alpha_1\geqslant\alpha_2\\ d_{\alpha_1}=d}} [\hat{Q}_{\alpha_1}(G), \hat{Q}_{\alpha_2}(G), \hat{Q}_{\beta}(G)]$$

$$\leqslant \prod_{\substack{\alpha_1+\alpha_2=\alpha\\ \alpha_1\geqslant\alpha_2\\ d_{\alpha_1}=d}} [\hat{Q}_{\alpha_2+\beta}(G), \hat{Q}_{\alpha_1}(G)]\,[\hat{Q}_{\alpha_1+\beta}(G), \hat{Q}_{\alpha_2}(G)].$$

Consider the factor $[\hat{Q}_{\alpha_2+\beta}(G), \hat{Q}_{\alpha_1}(G)]$. Now $\beta \leqslant \alpha_2 \leqslant \alpha_1$ but $d_{\alpha_2} \neq d_\beta$ so, by lemma 14·4 (ii),

$$\alpha_2+\beta+\alpha_1 = \alpha_1+\alpha_2+\beta = \alpha+\beta = \phi \quad \text{and} \quad d_{\alpha_2+\beta} = d_{\alpha_2} = d$$

so this factor is of the required form. The same argument applies to the other factor, completing this part of the proof.

(ii) The argument in this case is similar to that for part (i).

Finally, as an aid to determining the subgroups $Q_\phi(G)$, lemma 11·2 may be reduced to a more practical form for polyweights.

THEOREM 16·3. *Let ϕ be a function in Q and G be any group. Then*

$$Q_\phi(G) = \prod_{h=-1}^{d_\phi} \hat{Q}_{\phi_{(h)}}(G),$$

where, for each $h(-1 \leqslant h \leqslant d_\phi)$ the function $\phi_{(h)} \in Q$ is defined by

$$\begin{aligned} \phi_{(h)}(j) &= (\phi(h)+1)\,\delta_h(j) \quad && (j \leqslant h), \\ &= \phi(j) && (j > h). \end{aligned}$$

Proof. Notice first that $\phi_{(-1)} = \phi$ and $\phi_{(0)} = \phi+1$. By virtue of lemma 11·2 it is sufficient to prove that $\psi \geqslant \phi$ if and only if $\psi \gg \phi_{(h)}$ for some $h(-1 \leqslant h \leqslant d_\phi)$.

Suppose then that $\psi \geqslant \phi$. Then either $\psi = \phi$, in which case $\psi \gg \phi = \phi_{(-1)}$, or $\psi > \phi$, in which case there exists an integer $h \geqslant 0$ such that $\psi(h) > \phi(h)$ and $j > h \Rightarrow \psi(j) = \phi(j)$. Suppose first that $h > d_\phi$. Then $\psi(d_\phi) \geqslant k_{d_\phi+1} > \phi(d_\phi)$ and so by lemma 14·5 (iii) $\psi \gg (\phi(d_\phi)+1)\,\delta_{d_\phi} = \phi_{(d_\phi)}$. Now suppose that $h \leqslant d_\phi$. Then for $j > h$, $\psi(j) = \phi(j) = \phi_{(h)}(j)$ and for $j \leqslant h$, by lemma 14·5 (iii), $\psi(j) \gg \psi(h)\,\delta_h(j) \gg (\phi(h)+1)\,\delta_h(j) = \phi_{(h)}(j)$ and $\psi \gg \phi_{(h)}$.

Conversely, if $\psi \gg \phi_{(h)}$ for some h then $\psi \geqslant \phi_{(h)}$ and clearly $\phi_{(h)} \geqslant \phi$.

17. *Polyweight subgroups of a free group as intersections of commutator subgroups*

LEMMA 17·1. *Let $Q = Q^K$ be a polyweight range, c an integer $\geqslant 2$ and $c\delta_r = c\delta_r^K$ be the function given in definition 16·3. Then $\alpha+\beta = c\delta_r$ if and only if there exist positive integers m and n such that $\alpha = m\delta_r$ and $\beta = n\delta_r$, $m+n = c$ and either $m < k_{r+1}$ or $n < k_{r+1}$.*

Proof. Suppose first that $\alpha = m\delta_r$, $\beta = n\delta_r$, $m+n = c$ and one of m, n is $< k_{r+1}$. It may be assumed without loss of generality that $m \leqslant n$ so that $m < k_{r+1}$. The proof that $\alpha+\beta = c\delta_r$ is by induction over m. If $m = 1$ then $\alpha+\beta = n\delta_r+\delta_r = (n+1)\,\delta_r = c\delta_r$ by definition 14·3. If $1 < m < k_{r+1}$ then $\beta > (m-1)\,\delta_r \geqslant \delta_r$ and $d_\beta = r = d_{\delta_r}$ so that

$$\begin{aligned} \alpha+\beta &= \beta+\alpha \\ &= \alpha+(m-1)\,\delta_r+\delta_r \quad && \text{(by lemma 14·4 (ii))} \\ &= (m+n-1)\,\delta_r+\delta_r && \text{(by inductive hypothesis)} \\ &= c\delta_r. \end{aligned}$$

Now suppose that $\alpha+\beta = c\delta_r$; again it may be assumed without loss of generality that $\alpha \leqslant \beta$ so that $d_\alpha \leqslant d_\beta$. Write $m = \alpha(r)$ and $n = \beta(r)$. The definition of $\alpha+\beta$ then gives rise to two cases: case 1, $(\alpha+\beta)\,(j) = \alpha(j)+\beta(j)$ for all j and case 2, $d_\alpha = d_\beta = d$ say and $\alpha(d)+\beta(d) \geqslant k_{d+1}$.

Case 1, $(\alpha+\beta)(j) = \alpha(j)+\beta(j)$ for all j. Then $d_\beta = d_{\alpha+\beta} = d_{c\delta_r}$. Two subcases now arise: case 1·1, $c < k_{r+1}$ and case 1·2, $c \geqslant k_{r+1}$.

Case 1·1, $c < k_{r+1}$. Then $d_{c\delta_r} = r$ by lemma 14·5 (ii) so $d_\beta = r$. Then

$$m+n = (\alpha+\beta)(r) = (c\delta_r)(r) = c \quad \text{so that} \quad m < c < k_{r+1}.$$

If there exists $j \leqslant r$ such that $\alpha(j-1) > k_j\alpha(j)$ then

$$(c\delta_r)(j-1) = \alpha(j-1)+\beta(j-1) > k_j\alpha(j)+k_j\beta(j) = k_i(c\delta_r)(j)$$

which contradicts lemma 14·5 (ii): thus $\alpha(j-1) = k_j\alpha(j)$ for all $j \leqslant r$ so $\alpha = m\delta_r$. Similarly, $\beta = n\delta_r$.

Case 1·2, $c \geqslant k_{r+1}$. Then $d_\beta = d_{c\delta_r} = r+1$ and since $(c\delta_r)(r+1) = 1$, $\beta(r+1) = 1$ and $\alpha(r+1) = 0$. By the same argument as that given in case 1·1, $\alpha(j-1) = k_j\alpha(j)$ and $\beta(j-1) = k_j\beta(j)$ for all $j \leqslant r$. Then $d_\alpha = r$, since $\alpha(d_\alpha) > k_{d_\alpha+1}\alpha(d_\alpha+1)$ and $d_\alpha \leqslant r$, so $m = \alpha(r) < k_{r+1}$ and $\alpha = m\delta_r$. Also $n = \beta(r) \geqslant k_{r+1}$ since $d_\beta = r+1$ and then $\beta = n\delta_r$.

Case 2, $d_\alpha = d_\beta = d$ and $\alpha(d)+\beta(d) \geqslant k_{d+1}$. Then $d_{c\delta_r} = d_{\alpha+\beta} = d+1$ and $(c\delta_r)(d+1) = 1$. Thus, by lemma 14·5 (ii) and the fact that $c \geqslant 2$, $d = r$ and $c \geqslant k_{r+1}$. Thus $d_\alpha = d_\beta = r$ so that $m < k_{r+1}$ and $n < k_{r+1}$. Then, as before, $\alpha = m\delta_r$ and $\beta = n\delta_r$.

Corollary. *Let $Q = Q^K$ be a polyweight range, r be a positive integer and $\delta_r = \delta_r^K$. Then $\alpha+\beta = \delta_r$ if and only if there exist positive integers m and n such that $\alpha = m\delta_{r-1}$, $\beta = n\delta_{r-1}$ and $m+n = k_r$.*

Proof. This follows from the lemma since $\delta_r = k_{r-1}\delta_{r-1}$.

Lemma 17·2. *Let $Q = Q^K$ be a polyweight range and G be any group. Then for any integers $r \geqslant 0$ and $c \geqslant 1$*

$$\hat{Q}_{c\delta_r}(G) = \gamma_c(P_{K_r}(G)),$$

where $\delta_r = \delta_r^K$.

Proof. By induction over c. If $c = 1$ the result is given in theorem 16·1 (ii). Now suppose that $1 < c < k_{r+1}$. Then $d_{c\delta_r} = r$ and $(c\delta_r)(r) = c > 1$ so by theorem 16·2 (i)

$$\hat{Q}_{c\delta_r}(G) = \prod_{\substack{\alpha+\beta=c\delta_r\\ \alpha\geqslant\beta\\ d_\beta=r}} [\hat{Q}_\alpha(G), \hat{Q}_\beta(G)],$$

and by lemma 17·1 this may be written in the form

$$\hat{Q}_{c\delta_r}(G) = \prod_{\substack{m+n=c\\ m\geqslant n}} [\hat{Q}_{m\delta_r}(G), \hat{Q}_{n\delta_r}(G)].$$

Then, by the inductive hypothesis,

$$\hat{Q}_{c\delta_r}(G) = \prod_{\substack{m+n=c\\ m\geqslant n}} [\gamma_m(P_{K_r}(G)), \gamma_n(P_{K_r}(G))].$$

But every factor in this product is contained in $\gamma_{m+n}(P_{K_r}(G)) = \gamma_c(P_{K_r}(G))$ and one of the factors $(n=1)$ is $[\gamma_{c-1}(P_{K_r}(G)), P_{K_r}(G)] = \gamma_c(P_{K_r}(G))$ so $\hat{Q}_{c\delta_r}(G) = \gamma_c(P_{K_r}(G))$.

Finally suppose that $c \geqslant k_{r+1}$. Then $d_{c\delta_r} = r+1$ and $(c\delta_r)(r+1) = 1$ so by theorem 16·2 (ii),

$$\hat{Q}_{c\delta_r}(G) = \prod_{\substack{\alpha+\beta=c\delta_r\\ \alpha\geqslant\beta\\ d_\beta\ \ r}} [\hat{Q}_\alpha(G), \hat{Q}_\beta(G)]$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

and by lemma 17·1 this may be written in the form

$$\hat{Q}_{c\delta_r}(G) = \prod_{\substack{m+n=c\\ m\geqslant n\\ n<k_{r+1}}} [\hat{Q}_{m\delta_r}(G), \hat{Q}_{n\delta_r}(G)]$$
$$=\gamma_c(P_{K_r}(G))$$

as before.

Theorem. 17·1. *Let $Q = Q^K$ be a polyweight range, ϕ a function in Q and F an absolutely free group. Then*

$$\hat{Q}_{\phi}(F) = \bigcap_{r=0}^{d_\phi} \gamma_{\phi(r)}(P_{K_r}(F)).$$

Proof. By lemma 14·5 (iii),

$$\phi = \bigvee_{r=0}^{d_\phi} \phi(r)\,\delta_r$$

and by theorem 15·1 Q is partially collectable so lemma 13·3 may be applied repeatedly to yield

$$\hat{Q}_{\phi}(F) = \bigcap_{r=0}^{d_\phi} \hat{Q}_{\phi(r)\,\delta_r}(F).$$

The theorem now follows from lemma 17·2.

Comparison of the formulae of theorems 16·2 and 17·1 for particular examples of Q and ϕ yield non-trivial subgroup identities for absolutely free groups. One of these will be presented in detail as the next theorem to serve as an example.

Theorem 17·2. *For any absolutely free group F and positive integers m and n such that $m \leqslant n$,*

$$\gamma_{m+n}(F) \cap \delta(\gamma_m(F)) = [\gamma_m(F), \gamma_n(F)]\,[\gamma_{m+1}(F), \gamma_{n-1}(F)] \ldots [\gamma_n(F), \gamma_m(F)].$$

Proof. Let $K = (k_i)_{i=1}^{\infty}$ be the sequence $k_1 = m$, $k_j = 3$ $(j > 1)$ and $Q = Q^K$. Define the function $\phi: \omega \to \omega$ by $\quad \phi(0) = m+n, \quad \phi(1) = 2, \quad \phi(j) = 0 \quad (j > 1).$

Then clearly $\phi \in Q$. By theorem 16·2,

$$\hat{Q}_{\phi}(F) = [\gamma_m(F), \gamma_n(F)]\,[\gamma_{m+1}(F), \gamma_{n-1}(F)] \ldots [\gamma_n(F), \gamma_m(F)],$$

and by theorem 17·1, $\quad \hat{Q}_{\phi}(F) = \gamma_{m+n}(F) \cap \delta(\gamma_m(F)).$

Chapter IV. Centralizers

18. *The main lemma*

Definition 18·1. *Let A and B be normal subgroups of a group G. Then the* centralizer of A modulo B, $Z(A, B)$, *is defined by: $z \in Z(A, B)$ if and only if, for every $a \in A$, $[z, a] \in B$.*

Clearly then $Z(A, \{1\})$ is the centralizer of A and $Z(G, B)$ is the complete inverse image of the centre of G/B. The following relations are elementary

$$Z(A_1 A_2, B) = Z(A_1, B) \cap Z(A_2, B), \tag{1}$$

$$Z(A_1 \cap A_2, B) \supseteq Z(A_1, B)\,.\,Z(A_2, B), \tag{2}$$

$$Z(A, B_1 \cap B_2) = Z(A, B_1) \cap Z(A, B_2), \tag{3}$$

$$Z(A, B_1 B_2) \supseteq Z(A, B_1)\,.\,Z(A, B_2). \tag{4}$$

The object of this section is to prove a lemma which makes it easy to calculate the centralizers of shape subgroups modulo other ones within an absolutely free group.

For the remainder of this chapter it will be assumed that a fixed shape range W is being considered, in terms of which all definitions will be made.

DEFINITION 18·2. *For any pair* $\mathbf{a}$, $\mathbf{b}$ *of distinct basic commutators, the commutator* $[\mathbf{b}\leftarrow\mathbf{a}]$ *is defined recursively by*

(i) *if* $\mathbf{a} > \mathbf{b}$ *then* $[\mathbf{b}\leftarrow\mathbf{a}] = [\mathbf{a}\leftarrow\mathbf{b}]$,

(ii) *if* $\mathbf{a} < \mathbf{b}$, $\mathbf{b} = [\mathbf{b}_1, \mathbf{b}_2]$ *and* $\mathbf{a} < \mathbf{b}_2$ *then* $[\mathbf{b}\leftarrow\mathbf{a}] = [[\mathbf{b}_1\leftarrow\mathbf{a}], \mathbf{b}_2]$,

(iii) *otherwise* $[\mathbf{b}\leftarrow\mathbf{a}] = [\mathbf{b}, \mathbf{a}]$.

LEMMA 18·1. *With the notation of definition* 18·2,

(i) $[\mathbf{b}\leftarrow\mathbf{a}]$ *exists and is a basic commutator,*

(ii) $\sigma([\mathbf{b}\leftarrow\mathbf{a}]) = \sigma(\mathbf{b})+\sigma(\mathbf{a})$,

(iii) $\mathbf{b} < [\mathbf{b}\leftarrow\mathbf{a}]$ *and* $[\mathbf{b}\leftarrow\mathbf{a}] <^0 \mathbf{b}$ (*see definition* 8·3),

(iv) $[\mathbf{b}\leftarrow\mathbf{a}] \leqslant [\mathbf{b}, \mathbf{a}]$ *and* $[\mathbf{b}\leftarrow\mathbf{a}] \leqslant^0 [\mathbf{b}, \mathbf{a}]$,

(v) $\mathbf{b}_1 < \mathbf{b}_2 \Rightarrow [\mathbf{b}_1\leftarrow\mathbf{a}] < [\mathbf{b}_2\leftarrow\mathbf{a}]$ *and* $\mathbf{b}_1 <^0 \mathbf{b}_2 \Rightarrow [\mathbf{b}_1\leftarrow\mathbf{a}] <^0 [\mathbf{b}_2\leftarrow\mathbf{a}]$,

(vi) $[\mathbf{b}_1\leftarrow\mathbf{a}] = [\mathbf{b}_2\leftarrow\mathbf{a}] \Rightarrow \mathbf{b}_1 = \mathbf{b}_2$.

Proof. (i) and (ii) are proved together by induction over the commutator $[\mathbf{b}, \mathbf{a}]$. Suppose then that for all pairs $\mathbf{a}'$, $\mathbf{b}'$ of distinct basic commutators such that $[\mathbf{b}', \mathbf{a}'] < [\mathbf{b}, \mathbf{a}]$, $[\mathbf{b}'\leftarrow\mathbf{a}']$ exists and is a basic commutator and $\sigma([\mathbf{b}'\leftarrow\mathbf{a}']) = \sigma(\mathbf{b}')+\sigma(\mathbf{a}')$.

If $\mathbf{a} > \mathbf{b}$ then, by definition 6·1 (iii *c*), $[\mathbf{a}, \mathbf{b}] < [\mathbf{b}, \mathbf{a}]$ so by the inductive hypothesis $[\mathbf{b}\leftarrow\mathbf{a}] = [\mathbf{a}\leftarrow\mathbf{b}]$ exists and is a basic commutator and

$$\sigma[(\mathbf{b}\leftarrow\mathbf{a}]) = \sigma([\mathbf{a}\leftarrow\mathbf{b}]) = \sigma(\mathbf{b})+\sigma(\mathbf{a}).$$

If $\mathbf{a} < \mathbf{b}$, $\mathbf{b} = [\mathbf{b}_1, \mathbf{b}_2]$ and $\mathbf{b}_2 > \mathbf{a}$ then $[\mathbf{b}_1, \mathbf{a}] < [\mathbf{b}, \mathbf{a}]$ so $[\mathbf{b}_1\leftarrow\mathbf{a}]$ exists and is basic and $\sigma([\mathbf{b}_1, \mathbf{a}]) = \sigma(\mathbf{b}_1)+\sigma(\mathbf{a})$. Then

$$\begin{aligned}\sigma([\mathbf{b}\leftarrow\mathbf{a}]) &= \sigma([[\mathbf{b}_1\leftarrow\mathbf{a}], \mathbf{b}_2])\\ &= \sigma(\mathbf{b}_1)+\sigma(\mathbf{a})+\sigma(\mathbf{b}_2)\\ &= \sigma(\mathbf{b}_1)+\sigma(\mathbf{b}_2)+\sigma(\mathbf{a})\end{aligned}$$

(for $\mathbf{a} < \mathbf{b}_2 < \mathbf{b}_1$ since $\mathbf{b}$ is basic and so $\sigma(\mathbf{a}) \leqslant \sigma(\mathbf{b}_2) \leqslant \sigma(\mathbf{b}_1)$). It must be shown that if $[\mathbf{b}_1\leftarrow\mathbf{a}] = [\mathbf{c}_1, \mathbf{c}_2]$ then $\mathbf{c}_2 \leqslant \mathbf{b}_2$ (see definition 6·2 (A)). Now $\mathbf{a} < \mathbf{b}_2 < \mathbf{b}_1$ so there are only two possibilities: $\mathbf{b}_1 = [\mathbf{b}_{11}, \mathbf{b}_{12}]$ and $\mathbf{b}_{12} > \mathbf{a}$ or else $[\mathbf{b}_1\leftarrow\mathbf{a}] = [\mathbf{b}_1, \mathbf{a}]$ is basic. In the former case $[\mathbf{b}_1\leftarrow\mathbf{a}] = [[\mathbf{b}_{11}\leftarrow\mathbf{a}], \mathbf{b}_{12}]$ so $\mathbf{c}_2 = \mathbf{b}_{12} \leqslant \mathbf{b}_2$ since $\mathbf{b} = [\mathbf{b}_{11}, \mathbf{b}_{12}, \mathbf{b}_2]$ is basic. In the latter case, $\mathbf{c}_2 = \mathbf{b}_2 > \mathbf{a}$ by hypothesis.

Finally, if $[\mathbf{b}\leftarrow\mathbf{a}] = [\mathbf{b}, \mathbf{a}]$ is basic the result is immediately true.

(iii) follows immediately from (ii) since $\sigma(\mathbf{b}) < \sigma([\mathbf{b}\leftarrow\mathbf{a}])$.

(iv) This is proved by induction over $[\mathbf{b}, \mathbf{a}]$. Assume then that for any pair $\mathbf{a}', \mathbf{b}'$ of distinct basic commutators such that $[\mathbf{b}', \mathbf{a}'] < [\mathbf{b}, \mathbf{a}]$, $[\mathbf{b}'\leftarrow\mathbf{a}'] \leqslant^0 [\mathbf{b}', \mathbf{a}']$. Then there are three possibilities: if $\mathbf{a} > \mathbf{b}$ then $[\mathbf{b}\leftarrow\mathbf{a}] = [\mathbf{a}\leftarrow\mathbf{b}] \leqslant^0 [\mathbf{a}, \mathbf{b}] <^0 [\mathbf{b}, \mathbf{a}]$ by the inductive hypothesis. If $\mathbf{b} = [\mathbf{b}_1, \mathbf{b}_2]$ and $\mathbf{b}_2 > \mathbf{a}$ then $[\mathbf{b}\leftarrow\mathbf{a}] = [[\mathbf{b}_1\leftarrow\mathbf{a}], \mathbf{b}_2] \leqslant^0 [\mathbf{b}_1, \mathbf{a}, \mathbf{b}_2]$ since $[\mathbf{b}_1\leftarrow\mathbf{a}] \leqslant^0 [\mathbf{b}_1, \mathbf{a}]$ by the inductive hypothesis. But $\mathbf{b}$ is basic so $\mathbf{b}_1 > \mathbf{b}_2 > \mathbf{a}$ and thus $[\mathbf{b}_1, \mathbf{a}, \mathbf{b}_2] \leqslant^0 [\mathbf{b}, \mathbf{a}]$ by lemma 8·3. Finally, if $[\mathbf{b}\leftarrow\mathbf{a}] = [\mathbf{b}, \mathbf{a}]$ the result is immediate. The first statement, that $[\mathbf{b}\leftarrow\mathbf{a}] \leqslant [\mathbf{b}, \mathbf{a}]$, follows easily from this and part (ii) of the lemma.

(v) First observe that the second statement of this part follows from the first and part (ii). The first statement will be proved by induction over $\mathbf{b}_2$ and $\mathbf{a}$ ordered lexicographically, the inductive hypothesis being: suppose $\mathbf{a}'$, $\mathbf{b}_1'$ and $\mathbf{b}_2'$ are three basic commutators, $\mathbf{a}' \neq \mathbf{b}_1'$, $\mathbf{a}' \neq \mathbf{b}_2'$, $\mathbf{b}_1' < \mathbf{b}_2'$ and one of the conditions (*a*) $\mathbf{b}_2' < \mathbf{b}_2$ or (*b*) $\mathbf{b}_2' = \mathbf{b}_2$ and $\mathbf{a}' < \mathbf{a}$ holds, then $[\mathbf{b}_1' \leftarrow \mathbf{a}] < [\mathbf{b}_2' \leftarrow \mathbf{a}]$.

Now $\mathbf{b}_1 < \mathbf{b}_2$ so $\sigma(\mathbf{b}_1) \leqslant \sigma(\mathbf{b}_2)$ and thus, by part (ii),

$$\sigma([\mathbf{b}_1 \leftarrow \mathbf{a}] \leqslant \sigma([\mathbf{b}_2 \rightarrow \mathbf{a}]). \tag{5}$$

The definition of $[\mathbf{b}_2 \leftarrow \mathbf{a}]$ gives rise to three cases: case 1, $\mathbf{a} > \mathbf{b}_2$, case 2, $\mathbf{b}_2 = [\mathbf{b}_{21}, \mathbf{b}_{22}]$ and $\mathbf{b}_{22} > \mathbf{a}$ and case 3, $[\mathbf{b}_2, \mathbf{a}]$ is basic.

Case 1, $\mathbf{a} > \mathbf{b}_2$. Then $[\mathbf{b}_2 \leftarrow \mathbf{a}] = [\mathbf{a} \leftarrow \mathbf{b}_2]$ and there are two subcases: case 1·1, $\mathbf{a} = [\mathbf{a}_1, \mathbf{a}_2]$ and $\mathbf{a}_2 > \mathbf{b}_2$ and case 1·2, $[\mathbf{a}, \mathbf{b}_2]$ is basic.

Case 1·1, $\mathbf{a} = [\mathbf{a}_1, \mathbf{a}_2]$ and $\mathbf{a}_2 > \mathbf{b}_2$. Then

$$[\mathbf{b}_2 \leftarrow \mathbf{a}] = [\mathbf{a} \leftarrow \mathbf{b}_2] = [[\mathbf{a}_1 \leftarrow \mathbf{b}_2], \mathbf{a}_2]. \tag{6}$$

But $\mathbf{b}_2 > \mathbf{b}_1$ so $\mathbf{a} > \mathbf{a}_2 > \mathbf{b}_1$ also and thus

$$[\mathbf{b}_1 \leftarrow \mathbf{a}] = [\mathbf{a} \leftarrow \mathbf{b}_1] = [[\mathbf{a}_1 \leftarrow \mathbf{b}_1], \mathbf{a}_2]. \tag{7}$$

But $\mathbf{a}_1 < \mathbf{a}$ so by the inductive hypothesis, condition (*b*), $[\mathbf{b}_1 \leftarrow \mathbf{a}_1] < [\mathbf{b}_2 \leftarrow \mathbf{a}_1]$, that is,

$$[\mathbf{a}_1 \leftarrow \mathbf{b}_1] < [\mathbf{a}_1 \leftarrow \mathbf{b}_2]. \tag{8}$$

But $\sigma(\mathbf{a}_1) \geqslant \sigma(\mathbf{a}_2)$ since $\mathbf{a}$ is basic, so $\sigma([\mathbf{a}_1 \leftarrow \mathbf{b}_2]) > \sigma(\mathbf{a}_2)$ and thus, from proposition (6), $\mathrm{ld}([\mathbf{b}_2 \leftarrow \mathbf{a}]) = [\mathbf{a}_1 \leftarrow \mathbf{b}_2]$. Likewise, from proposition (7), $\mathrm{ld}([\mathbf{b}_1 \leftarrow \mathbf{a}]) = [\mathbf{a}_1 \leftarrow \mathbf{b}_1]$ and proposition (8) becomes $\mathrm{ld}([\mathbf{b}_1 \leftarrow \mathbf{a}]) < \mathrm{ld}([\mathbf{b}_2 \leftarrow \mathbf{a}])$. This, together with proposition (5) implies that $[\mathbf{b}_1 \leftarrow \mathbf{a}] < [\mathbf{b}_2 \leftarrow \mathbf{a}]$.

Case 1·2, $[\mathbf{a}, \mathbf{b}_2]$ is basic. Then

$$\begin{aligned}
[\mathbf{b}_1 \leftarrow \mathbf{a}] &\leqslant [\mathbf{b}_1, \mathbf{a}] && \text{by part (iv)} \\
&< [\mathbf{a}, \mathbf{b}_2] && \text{by lemma 8·1} \\
&= [\mathbf{b}_2 \leftarrow \mathbf{a}].
\end{aligned}$$

Case 2, $\mathbf{b}_2 = [\mathbf{b}_{21}, \mathbf{b}_{22}]$ and $\mathbf{b}_{22} > \mathbf{a}$. Then $[\mathbf{b}_2 \leftarrow \mathbf{a}] = [[\mathbf{b}_{21} \leftarrow \mathbf{a}], \mathbf{b}_{22}]$ and, since $\mathbf{b}_2$ is basic,

$$\mathrm{ld}([\mathbf{b}_2 \leftarrow \mathbf{a}]) = [\mathbf{b}_{21} \leftarrow \mathbf{a}]. \tag{9}$$

The definition of $[\mathbf{b}_1 \leftarrow \mathbf{a}]$ now gives rise to three subcases: case 2·1, $\mathbf{a} > \mathbf{b}_1$, case 2·2, $\mathbf{b}_1 = [\mathbf{b}_{11}, \mathbf{b}_{12}]$ and $\mathbf{b}_{12} > \mathbf{a}$ and case 2·3, $[\mathbf{b}_1, \mathbf{a}]$ is basic.

Case 2·1, $\mathbf{a} > \mathbf{b}_1$. Then $[\mathbf{b}_1 \leftarrow \mathbf{a}] = [\mathbf{a} \leftarrow \mathbf{b}_1]$ and there are two further subcases: case 2·1·1, $\mathbf{a} = [\mathbf{a}_1, \mathbf{a}_2]$ and $\mathbf{a}_2 > \mathbf{b}_1$ and case 2·1·2, $[\mathbf{a}, \mathbf{b}_1]$ is basic.

Case 2·1·1, $\mathbf{a} = [\mathbf{a}_1, \mathbf{a}_2]$ and $\mathbf{a}_2 > \mathbf{b}_1$. Then

$$[\mathbf{b}_1 \leftarrow \mathbf{a}] = [\mathbf{a} \leftarrow \mathbf{b}_1] = [[\mathbf{a}_1 \leftarrow \mathbf{b}_1], \mathbf{a}_2]. \tag{10}$$

By the assumption for case 2, $\mathbf{b}_{21} > \mathbf{b}_{22} > \mathbf{a}$ and, for case 2·1·1, $\mathbf{a}_2 > \mathbf{b}_1$. Thus

$$\mathbf{b}_{21} > \mathbf{a} > \mathbf{a}_2 > \mathbf{b}_1.$$

By the inductive hypothesis, condition (*a*), then

$$[\mathbf{b}_1 \leftarrow \mathbf{a}] < [\mathbf{b}_{21} \leftarrow \mathbf{a}]. \tag{11}$$

But $\mathbf{a}_1 < \mathbf{a}$ and, by the assumption for case 2, $\mathbf{a} < \mathbf{b}_2$. Thus by the inductive hypothesis, condition (*a*), again

$$[\mathbf{a}_1 \leftarrow \mathbf{b}_1] < [\mathbf{a} \leftarrow \mathbf{b}_1]. \tag{12}$$

Combining propositions (11) and (12) and using $[\mathbf{a} \leftarrow \mathbf{b}_1] = [\mathbf{b}_1 \leftarrow \mathbf{a}]$,

$$[\mathbf{a}_1 \leftarrow \mathbf{b}_1] < [\mathbf{b}_{21} \leftarrow \mathbf{a}]. \tag{13}$$

Since $\mathbf{a}$ is basic, $\mathbf{a}_1 > \mathbf{a}_2$ so $[\mathbf{a}_1 \leftarrow \mathbf{b}_1] > \mathbf{a}_2$ and then from proposition (10),

$$\mathrm{ld}\,([\mathbf{b}_1 \leftarrow \mathbf{a}]) = [\mathbf{a}_1 \leftarrow \mathbf{b}_1].$$

Combining this with propositions (9) and (13) yields

$$\mathrm{ld}\,([\mathbf{b}_1 \leftarrow \mathbf{a}]) < \mathrm{ld}\,([\mathbf{b}_2 \leftarrow \mathbf{a}])$$

and this, together with proposition (5), implies that $[\mathbf{b}_1 \leftarrow \mathbf{a}] < [\mathbf{b}_2 \leftarrow \mathbf{a}]$.

Case 2·1·2, $[\mathbf{a}, \mathbf{b}_1]$ is basic. In this case $[\mathbf{b}_1 \leftarrow \mathbf{a}] = [\mathbf{a}, \mathbf{b}_1]$ so

$$\mathrm{ld}\,([\mathbf{b}_1 \leftarrow \mathbf{a}]) = \mathbf{a} < [\mathbf{b}_{21} \leftarrow \mathbf{a}] = \mathrm{ld}\,([\mathbf{b}_2 \leftarrow \mathbf{a}])$$

and this, with proposition (5), implies that $[\mathbf{b}_1 \leftarrow \mathbf{a}] < [\mathbf{b}_2 \leftarrow \mathbf{a}]$.

Case 2·2, $\mathbf{b}_1 = [\mathbf{b}_{11}, \mathbf{b}_{12}]$ and $\mathbf{b}_{12} > \mathbf{a}$. Then $[\mathbf{b}_1 \leftarrow \mathbf{a}] = [[\mathbf{b}_{11} \leftarrow \mathbf{a}], \mathbf{b}_{12}]$ and so

$$\mathrm{ld}\,([\mathbf{b}_1 \leftarrow \mathbf{a}]) = [\mathbf{b}_{11} \leftarrow \mathbf{a}]. \tag{14}$$

Now $\mathbf{b}_1 = [\mathbf{b}_{11}, \mathbf{b}_{12}]$ so $\sigma(\mathbf{b}_1) > 1$ and $\mathbf{b}_1$ is basic so $\mathrm{ld}\,(\mathbf{b}_1) = \mathbf{b}_{11}$ and $\mathrm{tr}\,(\mathbf{b}_1) = \mathbf{b}_{12}$. Thus, since $\mathbf{b}_1 < \mathbf{b}_2$, definition 6·1 gives rise to three subcases: case 2·2·1, $\sigma(\mathbf{b}_1) < \sigma(\mathbf{b}_2)$, case 2·2·2, $\mathbf{b}_{11} < \mathbf{b}_{21}$ and case 2·2·3, $\mathbf{b}_{11} = \mathbf{b}_{21}$ and $\mathbf{b}_{12} < \mathbf{b}_{22}$.

Case 2·2·1, $\sigma(\mathbf{b}_1) < \sigma(\mathbf{b}_2)$. Then, by part (ii),

$$\sigma([\mathbf{b}_1 \leftarrow \mathbf{a}]) = \sigma(\mathbf{b}_1) + \sigma(\mathbf{a}) < \sigma(\mathbf{b}_2) + \sigma(\mathbf{a}) = \sigma([\mathbf{b}_2 \leftarrow \mathbf{a}])$$

so $[\mathbf{b}_1 \leftarrow \mathbf{a}] < \mathbf{b}_2 \leftarrow \mathbf{a}]$.

Case 2·2·2, $\mathbf{b}_{11} < \mathbf{b}_{21}$. Then by the inductive hypothesis, condition (*a*),

$$[\mathbf{b}_{11} \leftarrow \mathbf{a}] < [\mathbf{b}_{21} \leftarrow \mathbf{a}]$$

and this, together with propositions (4), (9) and (5), implies that $[\mathbf{b}_1 \leftarrow \mathbf{a}] < [\mathbf{b}_2 \leftarrow \mathbf{a}]$.

Case 2·2·3, $\mathbf{b}_{11} = \mathbf{b}_{21}$ and $\mathbf{b}_{12} < \mathbf{b}_{22}$. Then

$$\begin{aligned}[\mathbf{b}_1 \leftarrow \mathbf{a}] &= [[\mathbf{b}_{11} \leftarrow \mathbf{a}], \mathbf{b}_{12}] \\ &= [[\mathbf{b}_{21} \leftarrow \mathbf{a}], \mathbf{b}_{12}] \\ &< [[\mathbf{b}_{21} \leftarrow \mathbf{a}], \mathbf{b}_{22}] \\ &= [\mathbf{b}_2 \leftarrow \mathbf{a}].\end{aligned}$$

Case 2·3, $[\mathbf{b}_1, \mathbf{a}]$ is basic. Then $[\mathbf{b}_1 \leftarrow \mathbf{a}] = [\mathbf{b}_1, \mathbf{a}]$ so

$$\mathrm{ld}\,([\mathbf{b}_1 \leftarrow \mathbf{a}]) = \mathbf{b}_1. \tag{15}$$

Now $\mathbf{b}_1 < \mathbf{b}_2$ and definition 6·1 again yields three subcases: case 2·3·1, $\sigma(\mathbf{b}_1) < \sigma(\mathbf{b}_2)$, case 2·3·2, $\mathbf{b}_{11} < \mathbf{b}_{21}$ and case 2·3·3, $\mathbf{b}_{11} = \mathbf{b}_{21}$ and $\mathbf{b}_{12} < \mathbf{b}_{22}$.

Case 2·3·1 is the same as case 2·2·1.

Case 2·3·2, $\mathbf{b}_{11} < \mathbf{b}_{21}$. Then by the inductive hypothesis, condition (*a*),

$$[\mathbf{b}_{11} \leftarrow \mathbf{a}] < [\mathbf{b}_{21} \leftarrow \mathbf{a}], \tag{16}$$

and since $[\mathbf{b}_1, \mathbf{a}] = [\mathbf{b}_{11}, \mathbf{b}_{12}, \mathbf{a}]$ is basic, $\mathbf{b}_{12} \leqslant \mathbf{a}$. If $\mathbf{b}_{12} = \mathbf{a}$ then $[\mathbf{b}_{11} \leftarrow \mathbf{b}_{12}] = [\mathbf{b}_{11} \leftarrow \mathbf{a}]$ and if $\mathbf{b}_{12} < \mathbf{a}$ then by the inductive hypothesis, condition (*a*), again

$$[\mathbf{b}_{11} \leftarrow \mathbf{b}_{12}] = [\mathbf{b}_{12} \leftarrow \mathbf{b}_{11}] < [\mathbf{a} \leftarrow \mathbf{b}_{11}] = [\mathbf{b}_{11} \leftarrow \mathbf{a}]$$

so in either case $[\mathbf{b}_{11} \leftarrow \mathbf{b}_{12}] \leqslant [\mathbf{b}_{11} \leftarrow \mathbf{a}]$, which with proposition (16) gives

$$[\mathbf{b}_{11} \leftarrow \mathbf{b}_{12}] < [\mathbf{b}_{21} \leftarrow \mathbf{a}].$$

But $\mathbf{b}_1 = [\mathbf{b}_{11}, \mathbf{b}_{12}]$ is basic so $\mathbf{b}_1 = [\mathbf{b}_{11} \leftarrow \mathbf{b}_{12}] < [\mathbf{b}_{21} \leftarrow \mathbf{a}]$. This, together with propositions (15), (9) and (5), implies that $[\mathbf{b}_1 \leftarrow \mathbf{a}] < [\mathbf{b}_2 \leftarrow \mathbf{a}]$.

Case 2·3·3, $\mathbf{b}_{11} = \mathbf{b}_{21}$ and $\mathbf{b}_{12} < \mathbf{b}_{22}$. Now $[\mathbf{b}_1, \mathbf{a}]$ is basic so $\mathbf{b}_{12} \leqslant \mathbf{a}$. If $\mathbf{b}_{12} = \mathbf{a}$ then $[\mathbf{b}_{21} \leftarrow \mathbf{b}_{12}] = [\mathbf{b}_{21} \leftarrow \mathbf{a}]$ and if $\mathbf{b}_{12} < \mathbf{a}$ then by the inductive hypothesis, condition (*a*),

$$[\mathbf{b}_{21} \leftarrow \mathbf{b}_{12}] = [\mathbf{b}_{12} \leftarrow \mathbf{b}_{21}] < [\mathbf{a} \leftarrow \mathbf{b}_{21}] = [\mathbf{b}_{21} \leftarrow \mathbf{a}]$$

so in either case
$$[\mathbf{b}_{21} \leftarrow \mathbf{b}_{12}] \leqslant [\mathbf{b}_{21} \leftarrow \mathbf{a}]. \tag{17}$$

But ld $([\mathbf{b}_1 \leftarrow \mathbf{a}] = \mathbf{b}_1 = [\mathbf{b}_{11} \leftarrow \mathbf{b}_{12}] = [\mathbf{b}_{21} \leftarrow \mathbf{b}_{12}]$ which together with propositions (17) and (9) implies that ld $([\mathbf{b}_1 \leftarrow \mathbf{a}]) \leqslant$ ld $([\mathbf{b}_2 \leftarrow \mathbf{a}])$. Further tr $([\mathbf{b}_1 \leftarrow \mathbf{a}]) = \mathbf{a} < \mathbf{b}_{22}$ by the assumption for case 2 and $\mathbf{b}_{22} =$ tr $([\mathbf{b}_2 \leftarrow \mathbf{a}])$ so, using proposition (5), $[\mathbf{b}_1 \leftarrow \mathbf{a}] < [\mathbf{b}_2 \leftarrow \mathbf{a}]$ again.

Case 3, $[\mathbf{b}_2, \mathbf{a}]$ is basic. Then $[\mathbf{b}_1 \leftarrow \mathbf{a}] \leqslant [\mathbf{b}_1, \mathbf{a}] < [\mathbf{b}_2, \mathbf{a}] = [\mathbf{b}_2 \leftarrow \mathbf{a}]$.

(vi) follows immediately from (v).

Lemma 18·2. *Suppose* $\mathbf{x}$ *is an expression which is essentially* $<^0$ *some commutator* $\mathbf{c}$ (*see definition* 1·4) *and* $D\colon \mathbf{x} \rightarrow \mathbf{y}$. *Then* $\mathbf{y}$ *is also essentially* $<^0 \mathbf{c}$.

Proof. By checking the various parts of definition 8·1. For parts (i) to (xi), $\Xi(\mathbf{y}) \subseteq \Xi(\mathbf{x})$. For part (xii), $\Xi(\mathbf{x}) = \{[\mathbf{c}, \mathbf{b}, \mathbf{a}]\}$, $\Xi(\mathbf{y}) = \{[\mathbf{b}, \mathbf{a}, \mathbf{c}], [\mathbf{c}, \mathbf{a}, \mathbf{b}]\}$ and $[\mathbf{b}, \mathbf{a}, \mathbf{c}] <^0 \mathbf{x}$ and $[\mathbf{c}, \mathbf{a}, \mathbf{b}] <^0 \mathbf{x}$ by lemma 8·3. The remaining parts now follow easily.

Lemma 18·3. *Suppose* $\mathbf{a}$ and $\mathbf{b}$ *are two distinct commutators. Then* $D\colon [\mathbf{b}, \mathbf{a}] \rightarrow [\mathbf{b} \leftarrow \mathbf{a}]^{\epsilon} \mathbf{u}$ *where* $\mathbf{u}$ *is a* (*possibly empty*) *expression which is essentially* $<^0 [\mathbf{b} \leftarrow \mathbf{a}]$ *and* $\epsilon = \pm 1$.

Proof. (i) First it is shown that if $\mathbf{a}$ and $\mathbf{b}$ are basic commutators and $\mathbf{a} < \mathbf{b}$ then $D\colon [\mathbf{b}, \mathbf{a}] \rightarrow [\mathbf{b} \leftarrow \mathbf{a}]\, \mathbf{u}$ where $\mathbf{u}$ is a (possibly empty) expression which is essentially $<^0 [\mathbf{b} \leftarrow \mathbf{a}]$, this being done by induction over $[\mathbf{b}, \mathbf{a}]$. Either $[\mathbf{b}, \mathbf{a}]$ is basic, in which case $[\mathbf{b} \leftarrow \mathbf{a}] = [\mathbf{b}, \mathbf{a}]$ and $D\colon [\mathbf{b}, \mathbf{a}] \rightarrow [\mathbf{b} \leftarrow \mathbf{a}]$ trivially, or else $\mathbf{b} = [\mathbf{b}_1, \mathbf{b}_2]$ and $\mathbf{b}_2 > \mathbf{a}$ in which case $[\mathbf{b} \leftarrow \mathbf{a}] = [[\mathbf{b}_1 \leftarrow \mathbf{a}], \mathbf{b}_2]$. But then, by definition 8·1 (xii) and (i),

$$\begin{aligned} D\colon [\mathbf{b}, \mathbf{a}] &\rightarrow [\mathbf{b}_2, \mathbf{a}, \mathbf{b}_1]^{-1} [\mathbf{b}_1, \mathbf{a}, \mathbf{b}_2] \\ &\rightarrow [\mathbf{b}_1, \mathbf{a}, \mathbf{b}_2] [\mathbf{b}_2, \mathbf{a}, \mathbf{b}_1]^{-1}. \end{aligned} \tag{18}$$

By the inductive hypothesis since $\mathbf{b}_2 > \mathbf{a}$, $D\colon [\mathbf{b}_2, \mathbf{a}] \rightarrow [\mathbf{b}_2 \leftarrow \mathbf{a}]\, \mathbf{u}_1$ where $\mathbf{u}_1$ is essentially $<^0 [\mathbf{b}_2 \leftarrow \mathbf{a}]$. Then

$$\begin{aligned} D\colon [\mathbf{b}_2, \mathbf{a}, \mathbf{b}_1] &\rightarrow [[\mathbf{b}_2 \leftarrow \mathbf{a}]\, \mathbf{u}_1, \mathbf{b}_1] \\ &\rightarrow \mathbf{u}_2, \end{aligned} \tag{19}$$

where $$\mathbf{u}_2 = [[\mathbf{b}_2 \leftarrow \mathbf{a}], \mathbf{b}_1]\,[\mathbf{u}_1, \mathbf{b}_1] \tag{20}$$

and $$[\mathbf{u}_1, \mathbf{b}_1] \text{ is essentially } <^0 [[\mathbf{b}_2 \leftarrow \mathbf{a}], \mathbf{b}_1]. \tag{21}$$

Since $\mathbf{b}$ is basic, $\mathbf{b}_1 > \mathbf{b}_2 > \mathbf{a}$ so $\sigma(\mathbf{b}_1) \geqslant \sigma(\mathbf{b}_2) \geqslant \sigma(\mathbf{a})$ and then

$$\begin{aligned} \sigma([[\mathbf{b}_2 \leftarrow \mathbf{a}], \mathbf{b}_1]) &= \sigma(\mathbf{b}_2) + \sigma(\mathbf{a}) + \sigma(\mathbf{b}_1) \\ &\geqslant \sigma(\mathbf{b}_1) + \sigma(\mathbf{a}) + \sigma(\mathbf{b}_2) \\ &= \sigma([[\mathbf{b}_1 \leftarrow \mathbf{a}], \mathbf{b}_2]) \end{aligned} \tag{22}$$

by lemma 18·1 (ii) and definition 2·1 (v),

$$\operatorname{ld}([[\mathbf{b}_1 \leftarrow \mathbf{a}], \mathbf{b}_2]) = [\mathbf{b}_1 \leftarrow \mathbf{a}] \tag{23}$$

and $$\operatorname{ld}([[\mathbf{b}_2 \leftarrow \mathbf{a}], \mathbf{b}_1]) = [\mathbf{b}_2 \leftarrow \mathbf{a}] \quad \textit{or} \quad \mathbf{b}_1. \tag{24}$$

But $\mathbf{b}_1 > \mathbf{b}_2 > \mathbf{a}$ also implies, by lemma 18·1 (v), that $[\mathbf{b}_2 \leftarrow \mathbf{a}] < [\mathbf{b}_1 \leftarrow \mathbf{a}]$. Since $\mathbf{b}_1 < [\mathbf{b}_1 \leftarrow \mathbf{a}]$, this, with propositions (22), (23) and (24), shows that

$$[[\mathbf{b}_2 \leftarrow \mathbf{a}], \mathbf{b}_1] <^0 [[\mathbf{b}_1 \leftarrow \mathbf{a}], \mathbf{b}_2] = [\mathbf{b} \leftarrow \mathbf{a}],$$

so, using propositions (20) and (21), $\mathbf{u}_2$ is essentially $<^0$ $[\mathbf{b} \leftarrow \mathbf{a}]$. By propositions (18) and (19) and the inductive hypothesis,

$$\begin{aligned} D\colon [\mathbf{b}, \mathbf{a}] &\to [[\mathbf{b}_1 \leftarrow \mathbf{a}]\,\mathbf{u}_3, \mathbf{b}_2]\,\mathbf{u}_2 \\ &\to [[\mathbf{b}_1 \leftarrow \mathbf{a}], \mathbf{b}_2]\,[\mathbf{u}_3, \mathbf{b}_2]\,\mathbf{u}_2 \\ &= [\mathbf{b} \leftarrow \mathbf{a}]\,[\mathbf{u}_3, \mathbf{b}_2]\,\mathbf{u}_2, \end{aligned}$$

where $\mathbf{u}_3$ is essentially $<^0 [\mathbf{b}_1 \leftarrow \mathbf{a}]$. But then

$$[\mathbf{u}_3, \mathbf{b}_2] \text{ is essentially } <^0 [[\mathbf{b}_1 \leftarrow \mathbf{a}], \mathbf{b}_2] = [\mathbf{b} \leftarrow \mathbf{a}].$$

(ii) It remains to prove the lemma when $\mathbf{a} > \mathbf{b}$. In this case

$$D\colon [\mathbf{b}, \mathbf{a}] \to [\mathbf{a}, \mathbf{b}]^{-1} \to ([\mathbf{a} \leftarrow \mathbf{b}]\,\mathbf{u})^{-1}$$

by part (i), where $\mathbf{u}$ is essentially $<^0 [\mathbf{a} \leftarrow \mathbf{b}]$. But $[\mathbf{a} \leftarrow \mathbf{b}] = [\mathbf{b} \leftarrow \mathbf{a}]$ so

$$D\colon [\mathbf{b}, \mathbf{a}] \to \mathbf{u}^{-1}\,[\mathbf{b} \leftarrow \mathbf{a}]^{-1} \to [\mathbf{b} \leftarrow \mathbf{a}]^{-1}\,\mathbf{u}^{-1}.$$

LEMMA 18·4 (The main lemma). *Let* $\mathbf{x} = \mathbf{b}_1^{\beta_1}\mathbf{b}_2^{\beta_2}\ldots\mathbf{b}_k^{\beta_k}$ *be a basic expression other than* $\mathbf{1}$ *and let* $\mathbf{b}_r$ *be the maximum member, under the order* $<^0$, *of the set* $\Xi(\mathbf{x}) = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k\}$. *Let* $\mathbf{a}$ *be a basic commutator other than* $\mathbf{b}_r$. *Then*

$$D\colon [\mathbf{x}, \mathbf{a}] \to \mathbf{c}_1^{\gamma_1}\mathbf{c}_2^{\gamma_2}\ldots\mathbf{c}_l^{\gamma_l},$$

a basic expression other than $\mathbf{1}$, *and if* $\mathbf{c}_s$ *is the maximum member under* $<^0$ *of the set* $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_l$ *then* $\mathbf{c}_s = [\mathbf{b}_r \leftarrow \mathbf{a}]$ *and* $\gamma_s = \pm\beta_r$.

Proof. Using lemma 18·3, we have

$$\begin{aligned} D\colon [\mathbf{x}, \mathbf{a}] &\to [\mathbf{b}_1, \mathbf{a}]^{\beta_1}\,[\mathbf{b}_2, \mathbf{a}]^{\beta_2}\ldots[\mathbf{b}_k, \mathbf{a}]^{\beta_k} \\ &\to ([\mathbf{b}_1 \leftarrow \mathbf{a}]^{\epsilon_1}\mathbf{u}_1)^{\beta_1}\,([\mathbf{b}_2 \leftarrow \mathbf{a}]^{\epsilon_2}\mathbf{u}_2)^{\beta_2}\ldots([\mathbf{b}_k \leftarrow \mathbf{a}]^{\epsilon_k}\mathbf{u}_k)^{\beta_k} \\ &\to [\mathbf{b}_1 \leftarrow \mathbf{a}]^{\epsilon_1\beta_1}\mathbf{u}_1^{\beta_1}[\mathbf{b}_2 \leftarrow \mathbf{a}]^{\epsilon_2\beta_2}\mathbf{u}_2^{\beta_2}\ldots[\mathbf{b}_k \leftarrow \mathbf{a}]^{\epsilon_k\beta_k}\mathbf{u}_k^{\beta_k}, \end{aligned}$$

where, for each i, $\mathbf{u}_i$ is a (possibly empty) expression which is essentially $<^0 [\mathbf{b}_i \leftarrow \mathbf{a}]$ and hence by lemma 18·1 (v) essentially $<^0 [\mathbf{b}_r \leftarrow \mathbf{a}]$, each $\epsilon_i = \pm 1$ and for each $i \neq r$, $[\mathbf{b}_i \leftarrow \mathbf{a}] <^0 [\mathbf{b}_r \leftarrow \mathbf{a}]$. But then, for each $i \neq r$,

$$D\colon [\mathbf{b}_i \leftarrow \mathbf{a}]^{\epsilon_i \beta_i} \mathbf{u}_i^{\beta_i} \rightarrow \mathbf{y}_i,$$

where $\mathbf{y}_i$ is a product of basic commutators, all of which are $<^0 [\mathbf{b}_i \leftarrow \mathbf{a}]$ by lemma 18·2, and

$$D\colon \mathbf{u}_r^{\beta_r} \rightarrow \mathbf{z}$$

a product of basic commutators all of which are $<^0 [\mathbf{b}_r \leftarrow \mathbf{a}]$. Thus

$$D\colon [\mathbf{x}, \mathbf{a}] \rightarrow \mathbf{y}_1 \mathbf{y}_2 \ldots \mathbf{y}_{r-1} [\mathbf{b}_r \leftarrow \mathbf{a}]^{\epsilon_r \beta_r} \mathbf{z} \mathbf{y}_{r+1} \mathbf{y}_{r+2} \ldots \mathbf{y}_k,$$

which is a product of basic commutators all of which, apart from the factor $[\mathbf{b}_r \leftarrow \mathbf{a}]^{\epsilon_r \beta_r}$, are $<^0 [\mathbf{b}_r \leftarrow \mathbf{a}]$. The lemma follows.

19. *Applications*

Definition 19·1. *Let Φ and Ψ be two subsets of a shape range W. Then $\Psi - \Phi$ is the set of all minimal elements (under the coarse order) of the set*

$$\{\lambda\colon \Phi + \{\lambda\} \geqslant \Psi\}.$$

If α and β are elements of W and $\{\beta\} - \{\alpha\}$ contains exactly one element, then this element is denoted $\beta - \alpha$. The element $\beta \ominus \alpha$ is defined to be the minimum element (under the fine order) of the set

$$\{\lambda\colon \alpha + \lambda \geqslant \beta\}.$$

From this definition and definition 2·2 it follows immediately that $\Psi - \Phi$ may be defined by its properties:

(i) $\Phi + \Lambda \geqslant \Psi$ if an only if $\Lambda \geqslant \Psi - \Phi$ and

(ii) $\Psi - \Phi$ is totally unordered.

The element $\beta \ominus \alpha$ may be defined by its property that $\alpha + \lambda \geqslant \beta$ if and only if $\lambda \geqslant \beta \ominus \alpha$ and, provided the element $\beta - \alpha$ exists, it may be defined by its property that $\alpha + \lambda \geqslant \beta$ if and only if $\lambda \geqslant \beta - \alpha$.

Theorem 19·1. *Let Φ and Ψ be subsets of a partially collectable shape range W. Then, if F is an absolutely free group of rank at least* 3,

$$Z(\hat{W}_\Phi(F), \hat{W}_\Psi(F)) = \hat{W}_{\Psi - \Phi}(F).$$

Proof. Let $\rho\colon \mathbf{A} \rightarrow F$ be a free description of F. Suppose $z \in \hat{W}_{\Psi-\Phi}(F)$ and $\alpha \in \hat{W}_\Phi(F)$. Then there exist $\mathbf{z}, \mathbf{a} \in \mathbf{A}$ such that

$$\mathbf{z}\rho = z, \quad \mathbf{a}\rho = \mathbf{a}, \quad \Sigma(\mathbf{z}) \geqslant \Psi - \Phi \quad \text{and} \quad \Sigma(\mathbf{a}) \geqslant \Phi.$$

Then $\Sigma([\mathbf{z}, \mathbf{a}]) \geqslant \Psi$ and so $[z, a] \in \hat{W}_\Psi(F)$. This proves that $\hat{W}_{\Psi-\Phi}(F) \leqslant Z(\hat{W}_\Phi(F), \hat{W}_\Psi(F))$.

Now suppose $z \notin \hat{W}_{\Psi-\Phi}(F)$. Then, since W is partially collectable, there exists a non-negative integer c and a basic expression $\mathbf{z} = \mathbf{b}_1^{\beta_1} \mathbf{b}_2^{\beta_2} \ldots \mathbf{b}_k^{\beta_k} \in \mathbf{B}_{(c)} \cap \hat{\mathbf{B}}_{\Psi-\Phi}$ other than $\mathbf{1}$ such that $\mathbf{z}\rho = z$ modulo $\gamma_{c+1}(F) \,.\, \hat{W}_{\Psi+\Phi}(F)$. It may be assumed that c has been chosen as small as possible so that wt $(\mathbf{z}) = c$. Let $\mathbf{b}_r$ be the maximum member, under the order $<^0$, of the set $\Xi(\mathbf{z}) = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k\}$. Since $\mathbf{z} \in \hat{\mathbf{B}}_{\Psi-\Phi}$, $\Sigma(\mathbf{b}_r) \ngeqslant \Psi - \Phi$ and so $\Phi + \Sigma(\mathbf{b}_r) \ngeqslant \Psi$.

Thus there exists $\phi \in \Phi$ such that $\{\phi\}+\Sigma(\mathbf{b}_r) \not\geqslant \Psi$. But F is of rank at least 3, so by lemma 12·3 there exists a basic commutator $\mathbf{a}$ of shape exactly ϕ in $\mathbf{A}$ which is distinct from $\mathbf{b}_r$. Then, by lemma 18·4,

$$D\colon [\mathbf{z},\mathbf{a}] \to \mathbf{y} = \mathbf{c}_1^{\gamma_1}\mathbf{c}_2^{\gamma_2}\ldots\mathbf{c}_l^{\gamma_l},$$

a basic expression other than $\mathbf{1}$, and there exists an integer s $(1 \leqslant s \leqslant l)$ such that

$$\mathbf{c}_s = [\mathbf{b}_r \leftarrow \mathbf{a}] \quad \text{and} \quad \gamma_s = \pm\beta_r \neq 0.$$

Let $\mathrm{wt}\,(\mathbf{a}) = d$. Then $\mathrm{wt}\,([\mathbf{z},\mathbf{a}]) = c+d$ so $\mathbf{y}\rho = [\mathbf{z},\mathbf{a}]\rho$ modulo $\gamma_{c+d+1}(F)$ by theorem 8·1, corollary 4, and thus $\mathbf{y}\rho = [z,a]$ modulo $\gamma_{c+d+1}(F)\,.\,\hat{W}_\Psi(F)$ where $a = \mathbf{a}\rho$. Let $\mathbf{y}'$ be the basic expression made from $\mathbf{y}$ by deleting all commutators of weight $\geqslant c+d+1$ or shape set $\geqslant \Psi$. Then $\mathbf{y}'\rho = [z,a]$ modulo $\gamma_{c+d+1}(F)\,.\,\hat{W}_\Psi(F)$ also and $\mathbf{y}' \in \mathbf{B}_{(c)} \cap \hat{\mathbf{B}}_\Psi$. But

$$\mathrm{wt}\,(\mathbf{c}_s) = \mathrm{wt}\,([\mathbf{b}_r \leftarrow \mathbf{a}]) = c+d$$

and $$\Sigma(\mathbf{c}_s) = \{\hat{\sigma}(\mathbf{c}_s)\} = \{\hat{\sigma}(\mathbf{b}_r)+\hat{\sigma}(\mathbf{a})\} = \{\phi\}+\Sigma(\mathbf{b}_r) \not\geqslant \Psi$$

so $\mathbf{y}'$ contains the factor $\mathbf{c}_s^{\gamma_s}$ and is thus non-trivial. By theorem 9·1 (D) (iii) then

$$\mathbf{y}'\rho \notin \gamma_{c+d+1}(F)\,.\,\hat{W}_\Psi(F)$$

and thus $[z,a] \notin \gamma_{c+d+1}(F)\,.\,\hat{W}_\Psi(F)$ so that $[z,a] \notin \hat{W}_\Psi(F)$. But $\hat{\sigma}(\mathbf{a}) = \phi \in \Phi$ so $a \in \hat{W}_\Phi(F)$ and thus $z \notin Z(\hat{W}_\Phi(F), \hat{W}_\Psi(F))$.

Corollary 1. *Let α and β be elements of a partially collectable shape range W and suppose that the element $\beta-\alpha$ exists. Then, if F is an absolutely free group of rank at least* 3,

$$Z(\hat{W}_\alpha(F), \hat{W}_\beta(F)) = \hat{W}_{\beta-\alpha}(F).$$

Proof. Substitute $\Phi = \{\alpha\}$ and $\Psi = \{\beta\}$ in the theorem.

Corollary 2. *Let α and β be elements of a partially collectable shape range W. Then, if F is an absolutely free group of rank at least* 3,

$$Z(W_\alpha(F), W_\beta(F)) = W_{\beta\ominus\alpha}(F).$$

Proof. Translate corollary 1 according to the metatheorem of § 2.

Lemma 19·1. *Let $Q = Q^K$ be a polyweight range. Then, for any two elements α and β of Q, the element $\beta-\alpha$ exists and may be computed as follows*:

(i) *if $\alpha \geqslant \beta$ then $\beta-\alpha = 1$,*

(ii) *if $\alpha \not\geqslant \beta = \infty$ then $\beta-\alpha = \infty$ and*

(iii) *if $\alpha \not\geqslant \beta \neq \infty$, define a function ν from the non-negative integers to the integers by*

$$\begin{aligned} \nu(j) &= \beta(j)-\alpha(j) \qquad (\beta(j) \neq 1) \\ &= 0 \qquad\qquad\qquad (\beta(j) = 1). \end{aligned}$$

Then $$\beta-\alpha = \bigvee_{\nu(r)>0} \nu(r)\,\delta_r,$$

the join of the functions $\nu(r)\,\delta_r$ as given in definition 14·3 (ii) *for those values of r for which $\nu(r) > 0$.*

Proof. The definition of $\beta-\alpha$ given in parts (i) and (ii) of the lemma is obviously correct. It must now be shown that the definition given in part (iii) makes sense—that $\beta-\alpha$ is not being defined as the join of an empty collection of functions. To do this it is sufficient to show that there exists $r \in \omega$ such that $\nu(r) > 0$. Since $\alpha \not\geqslant \beta$, there exists $j \in \omega$ such that $\alpha(j) < \beta(j)$.

If $\beta(j) \neq 1$ then $r = j$ is the required integer, otherwise $\beta(j) = 1$ and $\alpha(j) = 0$ so

$$\beta(j-1) \geqslant k_j > \alpha(j-1) \quad \text{and} \quad r = j-1$$

is the required integer.

To show that the definition of $\beta - \alpha$ given in part (iii) is correct, it is sufficient to show that with this definition

$$\alpha + \xi \geqslant \beta \Leftrightarrow \xi \geqslant \beta - \alpha.$$

Suppose then that $\xi \geqslant \beta - \alpha$. Then for each $j \in \omega$,

$$\begin{aligned}(\alpha+\xi)(j) &\geqslant \alpha(j) + \xi(j)\\ &\geqslant \alpha(j) + (\beta-\alpha)(j)\\ &= \alpha(j) + \max_{\nu(r)>0}\{\nu(r)\,\delta_r(j)\}\\ &\geqslant \alpha(j) + \nu(j)\,\delta_j(j)\\ &= \alpha(j) + \nu(j).\end{aligned}$$

If $\beta(j) \neq 1$ then $\nu(j) = \beta(j) - \alpha(j)$ so $(\alpha+\xi)(j) \geqslant \alpha(j) + \nu(j) = \beta(j)$. If $\beta(j) = 1$ then $(\alpha+\xi)(j-1) \geqslant \beta(j-1) \geqslant k_j$ and then $(\alpha+\xi)(j) = 1 = \beta(j)$, so in either case

$$(\alpha+\xi)(j) \geqslant \beta(j).$$

Thus $\alpha + \xi \geqslant \beta$.

Conversely, suppose that $\alpha + \xi \geqslant \beta$. Then $(\alpha+\xi)(j) \geqslant \beta(j)$ for all $j \in \omega$. Then

$$\alpha(j) + \xi(j) = (\alpha+\xi)(j) \geqslant \beta(j)$$

so that $\xi(j) \geqslant \beta(j) - \alpha(j) \geqslant \nu(j)$ unless $d_\alpha = d_\xi = j-1$ and $\alpha(j-1) + \xi(j-1) \geqslant k_j$ in which case $(\alpha+\xi)(j) = 1$ so that $\beta(j) = 0$ or 1. Then $\nu(j) \leqslant 0$ so in either case $\xi(j) \geqslant \nu(j)$. Then

$$\nu(j) > 0 \Rightarrow \xi(j) > 0 \Rightarrow j \leqslant d_\xi$$

so by lemma 14·5 (iii),

$$\begin{aligned}\xi &= \bigvee_{j \leqslant d_\xi} \xi(j)\,\delta_j\\ &\geqslant \bigvee_{\nu(j)>0} \nu(j)\,\delta_j\\ &= \beta - \alpha.\end{aligned}$$

LEMMA 19·2. *Let W be a shape range such that $\beta - \alpha$ exists for every pair $\alpha, \beta \in \omega$. Then*

(i) *For any subset Ψ and element ϕ of W, $\Psi - \{\phi\}$ is equivalent (under the pre-order $\leqslant$) to the set $\{\psi - \phi : \psi \in \Psi\}$.*

(ii) *For any subsets Φ and Ψ of W, $\Psi - \Phi$ is equivalent (under the pre-order $\leqslant$) to the set*

$$\bigvee\{\Psi - \{\phi\} : \phi \in \Phi\}.$$

Proof. (i) Let $\{\lambda\} \geqslant \Psi - \{\phi\}$. Then $\{\phi + \lambda\} = \{\phi\} + \{\lambda\} \geqslant \Psi$ so there exists $\psi \in \Psi$ such that $\phi + \lambda \geqslant \psi$. Then $\lambda \geqslant \psi - \phi$ so $\{\lambda\} \geqslant \{\Psi - \phi : \psi \in \Psi\}$: but this argument is reversible, and thus

$$\{\lambda\} \geqslant \Psi - \{\phi\} \Leftrightarrow \{\lambda\} \geqslant \{\psi - \phi : \psi \in \Psi\}.$$

(ii) Let $\{\lambda\} \geqslant \Psi - \Phi$. Then $\Phi + \{\lambda\} \geqslant \Psi$ so, for any $\phi \in \Phi$, $\{\phi\} + \{\lambda\} \geqslant \Psi$ and thus $\{\lambda\} \geqslant \Psi - \{\phi\}$. Since this is true for any $\phi \in \Phi$, $\{\lambda\} \geqslant \bigvee\{\Psi - \{\phi\} : \phi \in \Phi\}$. Conversely, suppose that $\{\lambda\} \geqslant \bigvee\{\Psi - \{\phi\} : \phi \in \Phi\}$. Then, for any $\phi \in \Phi$, $\{\lambda\} \geqslant \Psi - \{\phi\}$ so $\{\lambda + \phi\} = \{\lambda\} + \{\phi\} \geqslant \Psi$. But, since this is true for any $\phi \in \Phi$, $\Phi + \{\lambda\} \geqslant \Psi$ and thus $\{\lambda\} \geqslant \Psi - \Phi$.

Finally two theorems are presented as examples of the application of this theory.

THEOREM 19·2. *Let* $G = F(\mathfrak{P}_{K_r} \wedge \mathfrak{N}_c)$ *where* $\mathfrak{P}_{K_r}$ *is a polynilpotent variety of length* $r \geqslant 2$, $c \geqslant 1$ *and the rank of the relatively free group* G *is at least* 3. *Then the centre of* G *is* $\gamma_c(G)$.

Proof. It is sufficient to prove that, if F is an absolutely free group of rank at least 3, then

$$Z(F, P_{K_r}(F) \,.\, \gamma_{c+1}(F)) = P_{K_r}(F) \,.\, \gamma_c(F).$$

Corresponding to the sequence K, form the polyweight range $Q = Q^K$. Then $\hat{Q}_{\delta_r}(F) = P_{K_r}(F)$ and $\hat{Q}_{\lambda_{c+1}}(F) = \gamma_{c+1}(F)$ by theorem 16·1 and the corollary to theorem 16·2. Thus $P_{K_r}(F) \,.\, \gamma_{c+1}(F) = \hat{Q}_\Phi(F)$ where $\Phi = \{\delta_r, \lambda_{c+1}\}$. By theorem 19·1 and lemma 19·2, since $F = \hat{W}_{\{1\}}(F)$, it is now sufficient to show that $\delta_r - 1 = \delta_r$ and $\lambda_{c+1} - 1 = \lambda_c$.

Since $r \geqslant 2$, $1(r-1) = 0$ so, defining ν as in lemma 19·1, $\nu(r-1) = \delta_r(r-1) = k_r$. Thus

$$\delta_r - 1 = \bigvee_{\nu(j)>0} \nu(j)\, \delta_j \geqslant \nu(r-1)\, \delta_{r-1} = k_r \delta_{r-1} = \delta_r.$$

But by definition 19·1, $\delta_r - 1 \leqslant \delta_r$. Thus $\delta_r - 1 = \delta_r$.

Now $\lambda_{c+1}(0) = c+1$ and $\lambda_{c+1}(1) = 1$ so, redefining ν to calculate $\lambda_{c+1} - 1$, $\nu(0) = c$ and $\nu(1) = 0$. Thus $\lambda_{c+1} - 1 = c\delta_0 = \lambda_c$.

COROLLARY 1. *With the hypotheses of the theorem, the upper and lower central series of* G *coincide.*

COROLLARY 2. *Let* $G = F(\mathfrak{P}_{K_r})$ *where* $\mathfrak{P}_{K_r}$ *is a polynilpotent variety of length* $r \geqslant 2$ *and the rank of the relatively free group* G *is at least* 3. *Then the centre of* G *is trivial.*

Proof. This can be proved by a simplified form of the argument used to prove the theorem or alternatively inferred directly from the theorem and the fact that G is residually nilpotent.

THEOREM 19·3. *Let* $G = F(\mathfrak{P}_{K_r} \vee \mathfrak{N}_c)$ *where* $\mathfrak{P}_{K_r}$ *is a polynilpotent variety of length* $r \geqslant 2$, $c \geqslant 1$ *and the rank of the relatively free group* G *is at least* 3. *Then the centre of* G *is* $P_{K_r}(G) \cap \gamma_c(G)$.

Proof. It is sufficient to prove that, if F is an absolutely free group of rank at least 3, then

$$Z(F, P_{K_r}(F) \cap \gamma_{c+1}(F)) = P_{K_r}(F) \cap \gamma_c(F).$$

This follows immediately from theorem 19·2, corollary 2 and equation (3) of § 18.

It will be observed that the restrictions on r and c assumed in these theorems are no more than non-triviality conditions. If $r = 1$ then $\mathfrak{P}_{K_r}$ is a nilpotent variety.

APPENDIX I. NON-COLLECTABILITY

In this appendix the fact stated in § 13 is established: that if α and β are elements of a shape range W, β finely dominates α and F is an absolutely free group of rank at least 3, then there exists an element in $W_\alpha(F)$ which cannot be described by a basic expression modulo $W_\beta(F)$.

If $\mathbf{b}$ is a basic commutator other than $\mathfrak{g}_0$, definition 18·2 yields a simpler definition of $[\mathbf{b} \leftarrow \mathfrak{g}_0]$, still by recursion over b:

(i) $[\mathfrak{g}_i \leftarrow \mathfrak{g}_0] = [\mathfrak{g}_i, \mathfrak{g}_0]$ and

(ii) If $\mathbf{b} = [\mathbf{b}_1, \mathbf{b}_2]$ then $[\mathbf{b} \leftarrow \mathfrak{g}_0] = [[\mathbf{b}_1 \leftarrow \mathfrak{g}_0], \mathbf{b}_2]$.

It is now necessary to generalize this idea slightly in two ways.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

DEFINITION I·1 (A). *If* $\mathbf{b}$ *is a basic commutator other than* $\mathbf{g}_0$ *and* n *a non-negative integer, then* $[\mathbf{b} \leftarrow n\mathbf{g}_0]$ *is defined recursively over* n:

(i) $[\mathbf{b} \leftarrow 0\mathbf{g}_0] = \mathbf{b}$,

(ii) $[\mathbf{b} \leftarrow n\mathbf{g}_0] = [[\mathbf{b} \leftarrow (n-1)\mathbf{g}_0] \leftarrow \mathbf{g}_0] \quad (n > 0)$.

(B) *If* $\mathbf{b}$ *is a basic commutator other than* $\mathbf{g}_0$, *then the expression* $[\mathbf{b} \leftarrow \mathbf{g}_0^{-1}]$ *is defined recursively over* $\mathbf{b}$:

(i) $[\mathbf{g}_i \leftarrow \mathbf{g}_0^{-1}] = [\mathbf{g}_i, \mathbf{g}_0^{-1}]$,

(ii) $[[\mathbf{b}_1, \mathbf{b}_2] \leftarrow \mathbf{g}_0^{-1}] = [[\mathbf{b}_1 \leftarrow \mathbf{g}_0^{-1}], \mathbf{b}_2] \quad (\mathbf{b}_2 \neq \mathbf{g}_0)$
$= [\mathbf{b}_1, \mathbf{b}_2, \mathbf{g}_0^{-1}] \quad (\mathbf{b}_2 = \mathbf{g}_0)$.

The following properties are elementary

(i) $[\mathbf{b} \leftarrow n\mathbf{g}_0]$ may be defined alternatively by recursion over $\mathbf{b}$ by $[\mathbf{g}_i \leftarrow n\mathbf{g}_0] = [\mathbf{g}_i, n\mathbf{g}_0]$ (see definition 7·3) and $[[\mathbf{b}_1, \mathbf{b}_2] \leftarrow n\mathbf{g}_0] = [[\mathbf{b}_1 \leftarrow n\mathbf{g}_0], \mathbf{b}_2]$.

(ii) $[[\mathbf{b} \leftarrow m\mathbf{g}_0] \leftarrow n\mathbf{g}_0] = [\mathbf{b} \leftarrow (m+n)\mathbf{g}_0]$.

(iii) $\sigma([\mathbf{b} \leftarrow n\mathbf{g}_0]) = \sigma(\mathbf{b})\,(+1)^n$ and $\sigma([\mathbf{b} \leftarrow \mathbf{g}_0^{-1}]) = \sigma(\mathbf{b}) + 1$.

(iv) If $[\mathbf{b}, \mathbf{g}_0]$ is basic, then there exists a generator $\mathbf{g}_i \neq \mathbf{g}_0$ and an integer $n \geqslant 0$ such that $\mathbf{b} = [\mathbf{g}_i, n\mathbf{g}_0]$.

DEFINITION I·2

(i) *If* $\mathbf{a}$ *and* $\mathbf{b}$ *are commutators, write* $\mathbf{b} \overset{\mu}{\geqslant} \mathbf{a}$ *if, for all* $i \geqslant 1$, $\mu_i(\mathbf{b}) \geqslant \mu_i(\mathbf{a})$ (*see definition* 12·1) *and* $\mathbf{b} \overset{\mu}{>} \mathbf{a}$ *if moreover there exists* $r \geqslant 1$ *such that* $\mu_r(\mathbf{b}) > \mu_r(\mathbf{a})$.

(ii) *If* $\mathbf{a}$ *is a commutator and* $\mathbf{x}$ *is an expression, write* $\mathbf{x} \overset{\mu}{\geqslant} \mathbf{a}$ *if* $\mathbf{x}$ *is essentially* $\overset{\mu}{\geqslant} \mathbf{a}$ *and* $\mathbf{x} \overset{\mu}{>} \mathbf{a}$ *if* $\mathbf{x}$ *is essentially* $\overset{\mu}{>} \mathbf{a}$.

Notice that $\overset{\mu}{\geqslant}$ is a pre-order, $\overset{\mu}{>}$ is not the corresponding strict relation in part (ii), the numbers of times $\mathbf{a}$ and $\mathbf{b}$ mention $\mathbf{g}_0$ are irrelevant in part (i) and that part (ii) is a bona fide extension of part (i). The following facts are elementary:

(i) If $\mathbf{a}$ is a commutator then $\mathbf{a} \overset{\mu}{\geqslant} \mathbf{a}$.

(ii) If $\mathbf{a}$ is a basic commutator other than $\mathbf{g}_0$ and n is a non-negative integer then $[\mathbf{a} \leftarrow n\mathbf{g}_0] \overset{\mu}{\geqslant} \mathbf{a}$ and $[\mathbf{a} \leftarrow \mathbf{g}_0^{-1}] \overset{\mu}{\geqslant} \mathbf{a}$ but the corresponding strict relation holds in neither case.

(iii) $\mathbf{x}_1\mathbf{x}_2 \overset{\mu}{\geqslant} \mathbf{a}$ if and only if both $\mathbf{x}_1 \overset{\mu}{\geqslant} \mathbf{a}$ and $\mathbf{x}_2 \overset{\mu}{\geqslant} \mathbf{a}$ and $\mathbf{x}_1\mathbf{x}_2 \overset{\mu}{>} \mathbf{a}$ if and only if both both $\mathbf{x}_1 \overset{\mu}{>} \mathbf{a}$ and $\mathbf{x}_2 \overset{\mu}{>} \mathbf{a}$.

(iv) If $\mathbf{x}_1 \overset{\mu}{\geqslant} \mathbf{a}_1$ and $\mathbf{x}_2 \overset{\mu}{\geqslant} \mathbf{a}_2$ then $[\mathbf{x}_1, \mathbf{x}_2] \overset{\mu}{\geqslant} [\mathbf{a}_1, \mathbf{a}_2]$ and if, moreover, either $\mathbf{x}_1 \overset{\mu}{>} \mathbf{a}_1$ or $\mathbf{x}_2 \overset{\mu}{>} \mathbf{a}_2$ then $[\mathbf{x}_1, \mathbf{x}_2] \overset{\mu}{>} [\mathbf{a}_1, \mathbf{a}_2]$.

LEMMA I·1. *Suppose* $\mathbf{c}$ *is a basic commutator other than* $\mathbf{g}_0$ *and* n *is a non-negative integer. Let* $\rho\colon \mathbf{A} \to G$ *be any description of a group* G. *Then there exists a (possibly empty) expression* $\mathbf{z}$ *such that*

(i) $[[\mathbf{c} \leftarrow n\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]\,\rho = ([[\mathbf{c} \leftarrow (n+1)\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{-1}\,[\mathbf{c} \leftarrow (n+1)\mathbf{g}_0]^{-1}\mathbf{z})\,\rho$ *and*

(ii) $\mathbf{z} \overset{\mu}{>} \mathbf{c}$.

Proof. The following easily checked group identity will be used: if a and b are elements of a group then

$$[a, b^{-1}] = [a, b, b^{-1}]^{-1}\,[a, b]^{-1}.$$

If $\mathbf{c} = [\mathbf{g}_i, k\mathbf{g}_0]$ for some non-negative integer k, then $[[\mathbf{c} \leftarrow n\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}] = [\mathbf{c}, n\mathbf{g}_0, \mathbf{g}_0^{-1}]$ and, using the identity just mentioned, we have

$$\begin{aligned}[\mathbf{c}, n\mathbf{g}_0, \mathbf{g}_0^{-1}]\rho &= ([\mathbf{c}, (n+1)\mathbf{g}_0, \mathbf{g}_0^{-1}]^{-1}[\mathbf{c}, (n+1)\mathbf{g}_0]^{-1})\rho \\ &= ([[\mathbf{c} \leftarrow (n+1)\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{-1}[\mathbf{c} \leftarrow (n+1)\mathbf{g}_0]^{-1})\rho,\end{aligned}$$

which is of the required form with $\mathbf{z}$ empty.

The proof may now proceed by induction over the weight of $\mathbf{c}$, assuming that $\mathbf{c} = [\mathbf{c}_1, \mathbf{c}_2]$, $\mathbf{c}_2 \neq \mathbf{g}_0$ and the result is true for $\mathbf{c}_1$. Then

$$[[\mathbf{c} \leftarrow n\mathbf{g}_0] \leftarrow g_0^{-1}] = [[[\mathbf{c}_1 \leftarrow n\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}], \mathbf{c}_2].$$

But then, by the inductive hypothesis,

$$[[\mathbf{c}_1 \leftarrow n\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]\rho = (\mathbf{d}^{-1}\mathbf{e}^{-1}\mathbf{z}_1)\rho,$$

where $\mathbf{d} = [[\mathbf{c}_1 \leftarrow (n+1)\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]$, $\mathbf{e} = [\mathbf{c}_1 \leftarrow (n+1)\mathbf{g}_0]$ and $\mathbf{z}_1 \overset{\mu}{>} \mathbf{c}_1$. Thus

$$\begin{aligned}[[\mathbf{c} \leftarrow n\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]\rho &= [\mathbf{d}^{-1}\mathbf{e}^{-1}\mathbf{z}_1, \mathbf{c}_2]\rho \\ &= ([\mathbf{d}^{-1}\mathbf{e}^{-1}, \mathbf{c}_2]\mathbf{z}_2)\rho\end{aligned}$$

where $\mathbf{z}_2 = [\mathbf{d}^{-1}\mathbf{e}^{-1}, \mathbf{c}_2, \mathbf{z}_1][\mathbf{z}_1, \mathbf{c}_2]$. Then

$$[\mathbf{d}^{-1}\mathbf{e}^{-1}, \mathbf{c}_2, \mathbf{z}_1] \overset{\mu}{\geqslant} [\mathbf{z}_1, \mathbf{c}_2] \overset{\mu}{>} [\mathbf{c}_1, \mathbf{c}_2] = \mathbf{c}$$

so $\mathbf{z}_2 \overset{\mu}{>} \mathbf{c}$. Then

$$\begin{aligned}[[\mathbf{c} \leftarrow n\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]\rho &= ([\mathbf{d}^{-1}, \mathbf{c}_2][\mathbf{d}^{-1}, \mathbf{c}_2, \mathbf{e}^{-1}][\mathbf{e}^{-1}, \mathbf{c}_2]\mathbf{z}_2)\rho \\ &= ([\mathbf{d}^{-1}, \mathbf{c}_2][\mathbf{e}^{-1}, \mathbf{c}_2]\mathbf{z}_3)\rho\end{aligned}$$

where $$\mathbf{z}_3 = [\mathbf{d}^{-1}, \mathbf{c}_2, \mathbf{e}^{-1}][[\mathbf{d}^{-1}, \mathbf{c}_2, \mathbf{e}^{-1}], [\mathbf{e}^{-1}, \mathbf{c}_2]]\mathbf{z}_2.$$

But $[\mathbf{d}^{-1}, \mathbf{c}_2] \overset{\mu}{\geqslant} \mathbf{c}_2$. Further, since $\mathbf{c} = [\mathbf{c}_1, \mathbf{c}_2]$ is basic, $\mathbf{c}_1 \neq \mathbf{g}_0$ so there exists $r \geqslant 1$ such that $\mu_r(\mathbf{c}) > 0$. Then $\mu_r(\mathbf{d}) = \mu_r(\mathbf{c}) > 0$ so $\mu_r([\mathbf{d}^{-1}, \mathbf{c}_2]) > \mu_r(\mathbf{c}_2)$. Thus $[\mathbf{d}^{-1}, \mathbf{c}_2] \overset{\mu}{>} \mathbf{c}_2$. Also $\mathbf{e} = [\mathbf{c}_1 \leftarrow (n+1)\mathbf{g}_0] \overset{\mu}{\geqslant} \mathbf{c}_1$ so $[\mathbf{d}^{-1}, \mathbf{c}_2, \mathbf{e}^{-1}] \overset{\mu}{>} \mathbf{c}$. Thus $\mathbf{z}_3 \overset{\mu}{>} \mathbf{c}$. Finally

$$\begin{aligned}[[\mathbf{c} \leftarrow n\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]\rho &= ([\mathbf{d}, \mathbf{c}_2]^{-1}[\mathbf{c}_2, \mathbf{d}, \mathbf{d}^{-1}]][\mathbf{e}, \mathbf{c}_2]^{-1}[\mathbf{c}_2, \mathbf{e}, \mathbf{e}^{-1}]\mathbf{z}_3)\rho \\ &= ([\mathbf{d}, \mathbf{c}_2]^{-1}[\mathbf{e}, \mathbf{c}_2]^{-1}\mathbf{z}_4)\rho,\end{aligned}$$

where $$\mathbf{z}_4 = [\mathbf{c}_2, \mathbf{d}, \mathbf{d}^{-1}][[\mathbf{c}_2, \mathbf{d}, \mathbf{d}^{-1}], [\mathbf{e}, \mathbf{c}_2]^{-1}][\mathbf{c}_2, \mathbf{e}, \mathbf{e}^{-1}]\mathbf{z}_3$$

and as before $\mathbf{z}_4 \overset{\mu}{>} \mathbf{c}$. The lemma is now proved since

$$[\mathbf{d}, \mathbf{c}_2] = [[[\mathbf{c}_1 \leftarrow (n+1)\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}], \mathbf{c}_2] = [[\mathbf{c} \leftarrow (n+1)\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]$$

and $$[\mathbf{e}, \mathbf{c}_2] = [[\mathbf{c}_1 \leftarrow (n+1)\mathbf{g}_0], \mathbf{c}_2] = [\mathbf{c} \leftarrow (n+1)\mathbf{g}_0].$$

LEMMA I·2. *If* $\mathbf{c}$ *is a commutator,* $\mathbf{x}$ *an expression,* $\mathbf{x} \overset{\mu}{>} \mathbf{c}$ *and* $E: \mathbf{x} \to \mathbf{y}$ *then* $\mathbf{y} \overset{\mu}{>} \mathbf{c}$.

Proof. Check the various parts of definition 9·1.

LEMMA I·3. *Let* $\mathbf{c}$ *be a basic commutator other than* $\mathbf{g}_0$ *and let* $\rho: \mathbf{A} \to G$ *be any description of a group* G. *Then to each non-negative integer n there exists an expression*

$$\mathbf{x}_n = \mathbf{u}_n[[\mathbf{c} \leftarrow n\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{(-1)^n}\mathbf{v}_n,$$

either $\mathbf{u}_n$ *or* $\mathbf{v}_n$ *possibly being empty, such that*

(i) $\mathbf{x}_n\rho = [\mathbf{c} \leftarrow \mathbf{g}_0^{-1}]\rho$,

(ii) $\mathbf{u}_0$ *is empty and if* $n \geqslant 1$, $\mathbf{u}_n$ *is of the form*

$$\mathbf{u}_n = \mathbf{b}_1^{\beta_1}\mathbf{b}_2^{\beta_2}\dots\mathbf{b}_k^{\beta_k} \in \mathbf{B}_{\alpha_n},$$

where $k \geqslant n$ *and* $\alpha_n = \sigma(\mathbf{c})\ (+1)^{n+1}$, *and*

(iii) *if* $\mathbf{v}_n$ *is non-empty, then* $\sigma(\mathbf{v}_n) \geqslant \alpha_n$ *and* $\mathbf{v}_n \overset{\mu}{>} \mathbf{c}$.

Proof. By induction over n. If $n = 0$, $[[\mathbf{c} \leftarrow 0\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}] = [\mathbf{c} \leftarrow \mathbf{g}_0^{-1}]$ and the lemma is true with both $\mathbf{u}_0$ and $\mathbf{v}_0$ empty.

Now suppose the result is true as stated for n. The corresponding result is proved for $n+1$. By lemma I·1, there exists an expression $\mathbf{z}_1$ (possibly empty) such that

$$[[\mathbf{c} \leftarrow n\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]\rho = ([[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{-1}\,[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0]^{-1}\,\mathbf{z}_1)\,\rho$$

and $\mathbf{z}_1 \overset{\mu}{>} \mathbf{c}$.

Hence $[\mathbf{c} \leftarrow \mathbf{g}_0^{-1}]\rho = \mathbf{x}'\rho$ where

$$\mathbf{x}' = \mathbf{u}_n([[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{-1}\,[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0]^{-1}\,\mathbf{z}_1)^{(-1)^n}\,\mathbf{v}_n.$$

Suppose n is even. Then

$$\mathbf{x}' = \mathbf{u}_n[[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{-1}\,[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0]^{-1}\,\mathbf{z}_1\mathbf{v}_n$$

and $$E\colon \mathbf{x}' \to \mathbf{u}_n[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0]^{-1}\,[[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{-1}\,\mathbf{z}_2\mathbf{z}_1\mathbf{v}_n,$$

where $$\mathbf{z}_2 = [[[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{-1}, [\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0]^{-1}]$$

and then $\mathbf{z}_2 \overset{\mu}{>} \mathbf{c}$ and $\sigma(\mathbf{z}_2) \geqslant \alpha_{n+1}$. Then

$$E\colon \mathbf{x}' \to \mathbf{u}_n[[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{-1}\,\mathbf{z}_2\mathbf{z}_1\mathbf{v}_n[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0]^{-1}\,z_3,$$

where $$\mathbf{z}_3 = [[[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{-1}, \mathbf{z}_2\mathbf{z}_1\mathbf{v}_n],$$

and then $\mathbf{z}_3 \overset{\mu}{>} \mathbf{c}$ and $\sigma(\mathbf{z}_3) \geqslant \alpha_{n+1}$. But now, since $\sigma(\mathbf{z}_2\mathbf{z}_1\mathbf{v}_n) \geqslant \alpha_n$, $E\colon \mathbf{z}_2\mathbf{z}_1\mathbf{v}_n \to \mathbf{w}\mathbf{z}_4$ where either $\mathbf{w}$ or $\mathbf{z}_4$ may be empty, but when they are not, $\mathbf{w} \in \mathbf{B}_{\alpha_{n+1}}$, $\sigma(\mathbf{w}) \geqslant \alpha_n$ and $\sigma(\mathbf{z}_4) \geqslant \alpha_{n+1}$. Further $\mathbf{w} \overset{\mu}{>} \mathbf{c}$ and $\mathbf{z}_4 \overset{\mu}{>} \mathbf{c}$. Thus

$$E\colon \mathbf{x}' \to \mathbf{u}_n[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0]^{-1}\,\mathbf{w}\mathbf{z}_4[[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{-1}\mathbf{z}_3.$$

It is now shown that

$$E\colon \mathbf{x}' \to \mathbf{u}_n\mathbf{w}'\mathbf{z}_5\mathbf{z}_4[[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{-1}\mathbf{z}_3,$$

where $\mathbf{w}' \in \mathbf{B}^W_{\alpha_{n+1}}$ but $\mathbf{w}' \neq \mathbf{1}$, $\sigma(\mathbf{w}') \geqslant \alpha_n$ and $\mathbf{z}_5$ may be empty but when it is not $\sigma(\mathbf{z}_5) \geqslant \alpha_{n+1}$ and $\mathbf{z}_5 \overset{\mu}{>} \mathbf{c}$. If $\mathbf{w}$ is empty or $\mathbf{1}$, this is true with $\mathbf{w}' = [\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0]^{-1}$ and $\mathbf{z}_5$ empty. Otherwise, suppose $\mathbf{w} = \mathbf{a}_1^{\gamma_1}\mathbf{a}_2^{\gamma_2}\dots\mathbf{a}_h^{\gamma_h}$. For each i $(1 \leqslant i \leqslant h)$, $\mathbf{a}_i \overset{\mu}{>} \mathbf{c}$ by definition I·2 (ii) since $\mathbf{w} \overset{\mu}{>} \mathbf{c}$. Thus for each i there exists r such that $1 \leqslant r < \tau$ and $\mu_r(\mathbf{c}) < \mu_r(\mathbf{a}_i)$. But $\mu_r([\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0]) = \mu_r(\mathbf{c})$ and so $[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0] \neq \mathbf{a}_i$. Thus $[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0]$ is not the same as any of the commutators $\mathbf{a}_i$ $(1 \leqslant i \leqslant h)$. But $[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0]$ is a basic commutator so there exists an integer t $(0 \leqslant t \leqslant h)$ such that

$$\mathbf{w}' = \mathbf{a}_i^{\gamma_1}\mathbf{a}_2^{\gamma_2}\dots\mathbf{a}_t^{\gamma_t}[\mathbf{c} \leftarrow (n+1)\,\mathbf{g}_0]^{-1}\,\mathbf{a}_{t+1}^{\gamma_{t+1}}\dots\mathbf{a}_h^{\gamma_h}$$

is a basic expression. But $\mathbf{w} \in \mathbf{B}_{\alpha_{n+1}}$ so each $\mathbf{a}_i$ is a commutator of shape $< \mathbf{a}_{n+1}$. Also $\sigma([\mathbf{c} \leftarrow (n+1)\mathbf{g}_0]) = \alpha_n$ so $\mathbf{w}' \in \mathbf{B}_{\alpha_{n+1}}$. But $\mathbf{w}' \neq \mathbf{1}$ since it contains the non-trivial factor $[\mathbf{c} \leftarrow (n+1)\mathbf{g}_0]$. Again $\sigma(\mathbf{w}) \geqslant \alpha_n$ so each $\mathbf{a}_i$ is of shape $\geqslant \alpha_n$ and $\sigma([\mathbf{c} \leftarrow (n+1)\mathbf{g}_0]) = \alpha_n$ so $\sigma(\mathbf{w}') \geqslant \alpha_n$. Then

$$E\colon [\mathbf{c} \leftarrow (n+1)\mathbf{g}_0]^{-1}\mathbf{w} \to \mathbf{w}'\mathbf{z}_5,$$

where $\sigma(\mathbf{z}_5) \geqslant \sigma([\mathbf{c} \leftarrow (n+1)\mathbf{g}_0]^{-1}\mathbf{w}) + 1 \geqslant \alpha_n + 1$ and since $\mathbf{z}_5$ is a product of commutators of $[\mathbf{c} \leftarrow (n+1)\mathbf{g}_0]^{-1}$ with other basic commutators, $\mathbf{z}_5 \overset{\mu}{>} \mathbf{c}$.

Finally, writing $\mathbf{u}_{n+1} = \mathbf{u}_n\mathbf{w}'$ and

$$\mathbf{v}_{n+1} = \mathbf{z}_5\mathbf{z}_4[\mathbf{z}_5\mathbf{z}_4, [[\mathbf{c} \leftarrow (n+1)\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{-1}\mathbf{z}_3,$$

$$E\colon \mathbf{x}' \to \mathbf{u}_{n+1}[[\mathbf{c} \leftarrow (n+1)\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{-1}\mathbf{v}_{n+1}$$

and $\mathbf{u}_{n+1}$ and $\mathbf{v}_{n+1}$ are of the required form.

The argument when n is odd is a slightly simplified version of the one just given.

THEOREM I·1. *Suppose α and β are elements of a shape range W, β finely dominates α and F is an absolutely free group of rank at least 3. Then there exists an element in $W_\alpha(F)$ which cannot be described by a basic expression modulo $W_\beta(F)$.*

Proof. Let $\rho\colon \mathbf{A} \to F$ be a free description of F, so that the number of generators of $\mathbf{A}$ is at least 3. Then by lemma 12·3 there exists a basic commutator $\mathbf{c}$ of shape exactly α in $\mathbf{A}$. It is now shown that $[\mathbf{c} \leftarrow \mathbf{g}_0^{-1}]\rho$ cannot be described by a basic expression modulo $W_\beta(F)$, for suppose that it can: then there exists a basic expression

$$\mathbf{w} = \mathbf{a}_1^{\gamma_1}\mathbf{a}_2^{\gamma_2}\ldots\mathbf{a}_n^{\gamma_n},$$

such that $[\mathbf{c} \leftarrow \mathbf{g}_0^{-1}]\rho = \mathbf{w}\rho$ modulo $W_\beta(F)$. Then, by lemma I·3, there exists an expression

$$\mathbf{x}_{n+1} = \mathbf{u}_{n+1}[[\mathbf{c} \leftarrow (n+1)\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]^{(-1)^{n+1}}\mathbf{v}_{n+1},$$

such that $\mathbf{x}_{n+1}\rho = [\mathbf{c} \leftarrow \mathbf{g}_0^{-1}]\rho$ where $\mathbf{u}_{n+1}$ is a basic expression of the form

$$\mathbf{u}_{n+1} = \mathbf{b}_1^{\beta_1}\mathbf{b}_2^{\beta_2}\ldots\mathbf{b}_k^{\beta_k} \quad (k \geqslant n+1) \quad \text{and} \quad \sigma(\mathbf{v}_{n+1}) \geqslant \alpha_{n+1}.$$

Also

$$\sigma([[\mathbf{c} \leftarrow (n+1)\mathbf{g}_0] \leftarrow \mathbf{g}_0^{-1}]) = \alpha_{n+1},$$

so

$$\mathbf{x}_{n+1}\rho = \mathbf{u}_{n+1}\rho \text{ modulo } W_{\alpha_{n+1}}(F).$$

But β dominates α, so $\alpha_{n+1} \leqslant \beta$. Now there exists $h(0 \leqslant j \leqslant n)$ such that

$$\mathbf{w}' = \mathbf{a}_1^{\gamma_1}\mathbf{a}_2^{\gamma_2}\ldots\mathbf{a}_h^{\gamma_h} \in \mathbf{B}_{\alpha_{n+1}},$$

and

$$\sigma(\mathbf{a}_{h+1}^{\gamma_{h+1}}\mathbf{a}_{h+2}^{\gamma_{h+2}}\ldots\mathbf{a}_n^{\gamma_n}) \geqslant \alpha_{n+1}.$$

Then $\mathbf{w}'\rho = [\mathbf{c} \leftarrow \mathbf{g}_0^{-1}]\rho = \mathbf{u}_{n+1}\rho$ modulo $W_{\alpha_{n+1}}(F)$ and so, by the basis theorem, $\mathbf{w}' = \mathbf{u}_{n+1}$ and so $h = k$. But it has already been observed that $h \leqslant n$ and $k \geqslant n+1$: this is a contradiction.

APPENDIX II. TERMS AND SYMBOLS USED IN THE TEXT

Symbols in more or less common use

Logic

$\Rightarrow$ logical implication. $\Leftrightarrow$ logical equivalence.

Set theory

For any property $\mathscr{P}$ that the elements of a set A may have, $\{x\colon x \in A, \mathscr{P}(x)\}$ is the set of all elements of A for which $\mathscr{P}(x)$ is true. When the set A is clear from the context, $\{x\colon \mathscr{P}(x)\}$ may be written.

$a \in A$ means 'a is a member (element) of A'. Occasionally $a, b \in A$ is used as shorthand for '$a \in A$ and $b \in A$'. $a \notin A$ means 'a is not a member of A'.

$\{a\}$ is the set whose only member is a, $\{a, b\}$ the set whose only members are a and b, and so on.

$(x_i)_{i=1}^n$, $(x_i)_{i=1}^\infty$, $(x_i)_{i<\tau}$, etc., are finite and transfinite sequences: $\{x_i\}_{i=1}^n$, $\{x_i\}_{i=1}^\infty$, $\{x_i\}_{i<\tau}$ are the corresponding sets.

$A \cap B$ and $A \cup B$ are the intersection and union respectively of the sets A and B. The intersection of a family of sets is written $\bigcap_{i=1}^{\infty} A_i$, $\bigcap_{i<\tau} A_i$, $\bigcap_{i \in I} A_i$, etc., and the intersection of a set $\mathscr{A}$ of sets is written $\cap \mathscr{A}$. In a similar fashion the symbol $\cup$ is used for unions and $\wedge$ and $\vee$ for meets and joins in a complete lattice.

$A \subseteq B$ means 'A is a subset of B'. $A - B$ is the complement of B in A, the set $\{x: x \in A, x \notin B\}$. $\varnothing$ is the empty set.

$\phi: A \to B$ indicates that ϕ is a function mapping the set A into the set B. Exceptions are the notations $d: \mathbf{x} \to \mathbf{y}$, $D: \mathbf{x} \to \mathbf{y}$, $e: \mathbf{x} \to \mathbf{y}$ and $E: \mathbf{x} \to \mathbf{y}$ which are given special definitions in §§ 8 and 9.

$|A|$ is the (cardinal) number of members of A. ω is the set of non-negative integers and the first infinite ordinal.

$m|n$, for integers m and n, means 'm divides n'.

Group theory

With the exception of the underlying group of a Lie ring, groups are written multiplicatively. The identity is denoted 1. $[x, y]$ is the element $x^{-1}y^{-1}xy$.

$A \leqslant B$ means 'A is a subgroup of B'. If A is a normal subgroup of B, the factor group is denoted B/A.

$A \cong B$ means 'A is isomorphic with B'.

AB is the subgroup generated by the normal subgroups A and B.

The subgroup generated by a family of normal subgroups is denoted $\prod_{i=1}^{\infty} A_i$, $\prod_{i<\tau} A_i$, $\prod_{i \in I} A_i$, etc.

$[A, B]$ is the subgroup generated by all commutators $[a, b]$ where $a \in A$ and $b \in B$: see also definition 1·5.

$\gamma_c(G)$, for positive integers c, are terms of the lower central series of the group G, defined by $\gamma_1(G) = G$ and $\gamma_c(G) = [\gamma_{c-1}(G), G]$ for $c > 1$.

$\delta^n(G)$, for non-negative integers n, are terms of the derived series of G, defined by $\delta^0(G) = G$ and $\delta^n(G) = [\delta^{n-1}(G), \delta^{n-1}(G)]$ for $n > 0$.

$\zeta_n(G)$, for non-negative integers n, are terms of the upper central series of G, defined by $\zeta_0(G) = \{1\}$ and $\zeta_n(G)$ is the complete inverse image of the centre of $G/\zeta_{n-1}(G)$ for $n > 0$.

Varieties

The language and notation concerned with this topic follows Hanna Neumann (1967). In particular,

F_τ, for some cardinal τ, is an absolutely free group of rank τ. F usually denotes a free group of arbitrary rank.

Varieties themselves are distinguished by Gothic script, $\mathfrak{B}$, $\mathfrak{N}_c$, and so on. The intersection and join of the varieties $\mathfrak{A}$ and $\mathfrak{B}$ are denoted $\mathfrak{A} \wedge \mathfrak{B}$ and $\mathfrak{A} \vee \mathfrak{B}$ respectively. $F_\tau(\mathfrak{B})$ is the free group of rank τ of the variety $\mathfrak{B}$ (a $\mathfrak{B}$-free group).

$\mathfrak{N}_c$ always denotes the variety of all groups which are nilpotent of class c. The notation associated with polynilpotent groups and varieties is described precisely in definition 14·1.

Algebras

The word 'algebra' is to be construed in the sense of 'universal algebra' as in Cohn (1965). Language and notation concerned with algebras follow this work.

Symbols defined in the text

References, unless otherwise stated, are to definition numbers.

$\mathbf{A}$	1·1	θ_r	7·5
$\mathbf{B}_\alpha$, $\hat{\mathbf{B}}_\alpha$, $\hat{\mathbf{B}}_\Phi$, $\mathbf{B}_{(c)}$	6·2	μ	1·1
$\mathbf{C}$	1·4	$\mu_i(\mathbf{c})$	12·1
$d\colon \mathbf{x}\to\mathbf{y}$, $D\colon \mathbf{x}\to\mathbf{y}$	8·1	ν	1·1
d_ϕ	14·2	Ξ	1·4
$e\colon \mathbf{x}\to\mathbf{y}$, $E\colon \mathbf{x}\to\mathbf{y}$	9·1	π, $\hat{\pi}$, Π, etc.	14·2
$\mathbf{G} = \{\mathbf{g}_i\}_{i<\tau}$	1·1	ρ	1·6
$\mathcal{G}$	1·6	σ, $\hat{\sigma}$, Σ	2·3
ht	page 346	τ	1·1
$H^+(\phi)$, $H^-(\phi)$	16·1	χ	1·1
K, K_r	14·1	Ω	1·1
ld	6·1	1	2·1, 14·2
N^-	1·3	$\mathbf{1}$	1·1
$\mathbf{P}_r$, $\mathbf{P}_r(w)$	7·5	∞	1·3, 2·1, 14·2
$P_{K_r}(G)$, $\mathfrak{P}_{K_r}$	14·1	$\leqslant$, $<$, etc.	1·3, 2·1, 6·1, 14·2
Q, $\mathfrak{Q}$, $\mathbf{Q}_\phi$, etc.	14·2	$\preccurlyeq$, $\prec$, etc.	2·1, 2·2, 14·2
$S(\mathbf{p})$	7·5	$\leqslant^0$, $<^0$	8·3
tr	6·1	$\overset{\mu}{\geqslant}$, $\overset{\mu}{>}$	1·2
$U(\phi)$, $\hat{U}(\phi)$	5·1	$+$	2·1, 2·2, 14·2
$\hat{U}_1(\phi)$, $\hat{U}_2(\phi)$	16·1	$(+1)^n$	9·2
wt	1·3	$-$, $\ominus$	19·1
W	2·1	$\wedge$, $\vee$	2·1, 2·2, 14·2
$\mathbf{W}_\alpha$, $\hat{\mathbf{W}}_\alpha$, $\hat{\mathbf{W}}_\Phi$	3·1	$[\mathbf{b}, n_1\mathbf{a}_1, n_2\mathbf{a}_2, \ldots, n_k\mathbf{a}_k]$	7·3
$W_\alpha(G)$, $\hat{W}_\alpha(G)$, $\hat{W}_\Phi(G)$	3·2	$[\mathbf{b}\leftarrow\mathbf{a}]$	18·2
$\mathfrak{W}_\alpha$, $\hat{\mathfrak{W}}_\alpha$, $\hat{\mathfrak{W}}_\Phi$	3·2	$[\mathbf{b}\leftarrow n\mathbf{g}_0]$, $[\mathbf{b}\leftarrow \mathbf{g}_0^{-1}]$	1·1
$\mathbf{X}_i$	7·4	$[A, B]$	1·5
$Z(A, B)$	18·1	$\mathbf{b}^+$	7·2
δ_r, $n\delta_r$	14·3	ϕ_{+h}, ϕ_{-h}	16·1
ϵ	1·1		

Terms defined in the text

References

Baumslag, G. 1963 Wreath products and extensions. *Math. Z.* **81**, 286–299.

Cohn, P. M. 1965 *Universal algebra*. New York: Harper and Row.

Gorčakov, Yu. M. 1967 Multinil'potentnye gruppy. *Alg. i Log. Seminar* **6**, part 3, 25–30.

Hall, M. Jun. 1950 A basis for free Lie rings and higher commutators in free groups. *Proc. Am. Math. Soc.* **1**, 575–581.

Hall, P. 1934 A contribution to the theory of groups of prime-power order. *Proc. Lond. Math. Soc.* **36**, 29–95.

Magnus, W. 1935 Beziehungen zwischen Gruppen und Idealen in einem speziellen Ring. *Math. Annalen* **111**, 259–280.

Magnus, W. 1937 Über Beziehungen zwischen höheren Kommutatoren. *J. reine angew. Math.* **177**, 105–115.

Neumann, Hanna 1967 *Varieties of groups*. Berlin: Springer-Verlag.

Ramsey, F. P. 1928 On a problem in formal logic. *Proc. Lond. Math. Soc.* (2), **30**, 338–384.

Witt, E. 1937 Treue Darstellung Liescher Ringe. *J. reine angew. Math.* **177**, 152–160.